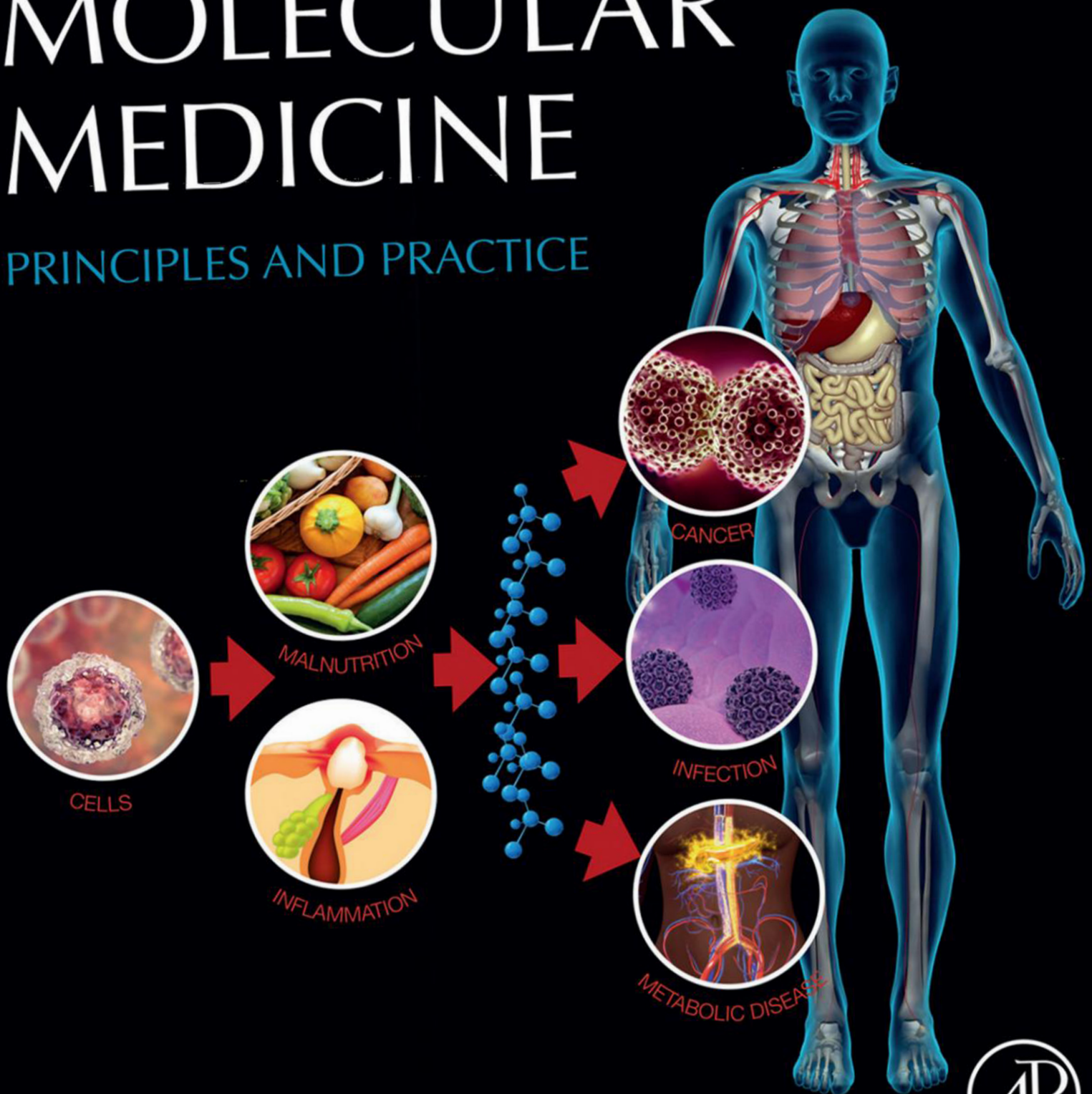


# CLINICAL MOLECULAR MEDICINE

PRINCIPLES AND PRACTICE



Edited by  
Dhavendra Kumar







# CLINICAL MOLECULAR MEDICINE

---



# CLINICAL MOLECULAR MEDICINE

## Principles and Practice

---

*Edited by*

DHAVENDRA KUMAR

*The William Harvey Research Institute, Barts and The London School of Medicine & Dentistry,  
Queen Mary University of London, London, United Kingdom*



ACADEMIC PRESS

An imprint of Elsevier

Academic Press is an imprint of Elsevier  
125 London Wall, London EC2Y 5AS, United Kingdom  
525 B Street, Suite 1650, San Diego, CA 92101, United States  
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States  
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Copyright © 2020 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: [www.elsevier.com/permissions](http://www.elsevier.com/permissions).

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

### Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

### British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

### Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-809356-6

For Information on all Academic Press publications  
visit our website at <https://www.elsevier.com/books-and-journals>

*Publisher:* Andre Gerhard Wolff  
*Acquisition Editor:* Erin Hill-Parks  
*Editorial Project Manager:* Megan Ashdown  
*Production Project Manager:* Stalin Viswanathan  
*Cover Designer:* Miles Hitchen

Typeset by MPS Limited, Chennai, India





# Dedication

---



Dedicated to The Late Sir David Weatherall (1933–2019) GBE, FRS, DM, FRCP

Emeritus Regius Professor of Medicine, University of Oxford, Oxford,  
United Kingdom

Founder/Director—Weatherall Institute of Molecular Medicine,  
University of Oxford, Oxford, United Kingdom



# Dedication

---

To my

Parents showering blessings from the heaven

Wife, Anju, for life-long support, confidence, and sharing my passion for books

Children, Ashish and Jaime, Nikita and Mayank for their goodwill and trusting me

Grand children, Jaya, Arya, Anya, and Krish, for their joyful presence and a reminder of the future





# Contents

---

<b>List of contributors</b>	<b>xv</b>		
<b>About the author</b>	<b>xix</b>		
<b>Foreword</b>	<b>xxi</b>		
<b>Preface</b>	<b>xxiii</b>		
<b>Acknowledgement and Disclaimer</b>	<b>xv</b>		
<b>I</b>			
<hr/>			
<b>FUNDAMENTALS OF MOLECULAR MEDICINE</b>			
<b>1. The human genome and molecular medicine</b>	<b>3</b>		
DHAVENDRA KUMAR			
1.1 Introduction	3		
1.2 Hereditary factors: genes, genetics, and genomics	3		
1.3 Human genome variation and human disease	7		
1.4 The mitochondrial genome	10		
1.5 Functional genomics, transcriptomics, and proteomics	12		
1.6 Translational human genomics	14		
1.7 Human genomics for socioeconomic development	14		
1.8 Conclusions	15		
References	15		
<b>2. Cellular structure and molecular cell biology</b>	<b>17</b>		
LUCIANA AMARAL HADDAD			
2.1 Plasma membrane	17		
2.2 Cytoskeleton	21		
2.3 Cell nucleus and gene expression	23		
2.4 Protein synthesis	32		
2.5 Vesicular trafficking: the secretory and endocytic pathways	34		
2.6 Protein turnover and cell size control	38		
2.7 Microtubule-organizing centers and the cell cycle	39		
2.8 Energy production and oxidative stress	42		
2.9 Summary	44		
Bibliography	45		
<b>3. Molecular basis of clinical metabolomics</b>	<b>47</b>		
DIEGO F. GOMEZ-CASATI AND MARÍA V. BUSI			
3.1 Introduction	47		
3.2 Clinical applications of metabolomics	47		
3.3 Techniques used in metabolomics and databases		51	
3.4 Conclusions		52	
Acknowledgments		52	
References		52	
<b>4. Clinical applications of next-generation sequencing</b>		<b>57</b>	
IAN TULLY			
4.1 Introduction		57	
4.2 Next-generation sequencing technologies		58	
4.3 Choice of test		59	
4.4 Whole-genome sequencing		62	
4.5 Summary		62	
4.6 Ethical considerations		62	
4.7 Bioinformatics		65	
4.8 Clinical interpretation of variants		66	
<b>II</b>			
<hr/>			
<b>MOLECULAR MEDICINE IN CLINICAL PRACTICE</b>			
<b>5. Molecular basis of obesity disorders</b>		<b>73</b>	
LOTTE KLEINENDORST AND MIEKE M. VAN HAELEST			
5.1 Introduction		73	
5.2 Clinical cases		74	
5.3 Molecular systems underpinning the clinical scenario		77	
5.4 Molecular biology and pathophysiology of obesity		79	
5.5 Targeted molecular diagnosis and therapy		83	
5.6 Summary		86	
References		86	
Guide to further reading: articles		87	
<b>6. Molecular dysmorphology</b>		<b>89</b>	
KOUMUDI GODBOLE			
6.1 Introduction		89	
6.2 Clinical cases and molecular basis		91	
6.3 Molecular diagnosis and therapy		98	
6.4 Conclusion/summary		103	
References		104	

<b>7. Disorders of sex development</b>	<b>107</b>	<b>11. Molecular pathophysiology of systemic hypertension</b>	<b>169</b>
PIERRE SINAJON, SALEH ALBANYAN AND DAVID CHITAYAT		TATIANA GAROFALIDOU AND PATRICIA B. MUNROE	
7.1 Introduction	107	11.1 Introduction	169
7.2 Sex chromosome disorder of sex development	108	11.2 Blood pressure regulation—key systems	169
7.3 46, XY disorders of sexual differentiation	110	11.3 Genetics of hypertension	170
7.4 46, XX disorders of sex development	113	11.4 Monogenic forms of systemic hypertension	171
7.5 Investigations	115	11.5 Familial hyperaldosteronism; type 1—glucocorticoid remediable aldosteronism	181
7.6 Gender assignment	116	11.6 Genetic overlap of monogenic and essential hypertension	183
7.7 Gonadal cancer risk	117	11.7 Clinical implications from genetic studies of hypertension	184
7.8 Conclusion	117	11.8 Future perspectives	184
7.9 Resources Web-based resources	117	Acknowledgments	185
References	117	References	185
<b>8. Molecular systems in cardiovascular developmental disorders</b>	<b>121</b>	<b>12. Molecular basis of stroke</b>	<b>189</b>
ALAN GRAHAM STUART		YOSHIJI YAMADA	
8.1 Introduction	121	12.1 Introduction	189
8.2 Historical overview	121	12.2 Genetics of stroke	189
8.3 Normal development of the heart	121	12.3 Single-gene disorders associated with stroke	189
8.4 Advances in technology	122	12.4 Genetics of common forms of stroke	191
8.5 Chromosomal aneuploidy and structural heart disease	123	12.5 Clinical implication	207
8.6 Single-gene (Mendelian) disorders	124	12.6 Conclusion	207
8.7 The noncoding regulatory genome in congenital heart disease: microRNAs and circular RNAs	126	References	208
8.8 Noncardiac congenital anomalies in congenital heart disease	128	Further reading	216
8.9 The epigenome and congenital heart disease	128	<b>13. Clinical molecular endocrinology</b>	<b>217</b>
8.10 Future considerations	129	ATUL KALHAN	
References	129	13.1 Introduction	217
<b>9. Channelopathies in clinical medicine—cardiac arrhythmias</b>	<b>133</b>	13.2 Congenital hypopituitarism, congenital hypogonadotropic hypogonadism, and pituitary adenoma	217
JULIAN O.M. ORMEROD AND ELIJAH R. BEHR		13.3 Primary hyperparathyroidism and multiple endocrine neoplasia syndromes	222
9.1 Introduction	133	13.4 Chronic hypocalcemia and hypophosphatemia	227
9.2 Clinical cases	133	13.5 Primary hyperaldosteronism	230
9.3 Molecular systems underpinning the clinical scenario	135	13.6 Congenital adrenal hyperplasia, apparent mineralocorticoid excess, and renal tubular defects	233
9.4 Overview of molecular biology and pathophysiology	137	13.7 Pheochromocytoma and paraganglioma	236
9.5 Short QT syndrome	144	13.8 Primary gonadal failure and androgen resistance/insensitivity syndrome	240
9.6 Catecholaminergic polymorphic ventricular tachycardia	145	References	242
9.7 Brugada syndrome	146	<b>14. Genetic disorders of lipoprotein metabolism</b>	<b>245</b>
9.8 Targeted molecular diagnosis and therapy	147	DEEPAK BHATNAGAR, JONATHAN SCHOFIELD AND HANDREAN SORAN	
9.9 Summary	150	14.1 Introduction	245
References	150	14.2 An outline of lipoprotein metabolism	245
<b>10. Chronic heart failure</b>	<b>153</b>	14.3 The environmental and genetic factors affecting lipid metabolism	247
ABBAS ZAIDI AND PARIN SHAH		14.4 Screening for lipoprotein disorders	249
10.1 Introduction	153		
10.2 Epidemiology	153		
10.3 Etiology of heart failure	154		
10.4 Clinical assessment	155		
10.5 Genetic testing	161		
10.6 Management	164		
References	166		

14.5 Diagnosing genetic disorders of lipoprotein metabolism	249	17.6 Human leukocyte antigen—antibody-detection techniques	319
14.6 Common (polygenic) hypercholesterolemia	249	17.7 Human leukocyte antigen and blood transfusion	322
14.7 Familial hypercholesterolemia	252	17.8 Conclusions	323
14.8 Characteristic clinical features of familial hypercholesterolemia	252	References	323
14.9 Screening strategies for familial hypercholesterolemia	255	Further reading	325
14.10 Genetic disorders resulting in hypertriglyceridemia	255	<b>18. Disorders of abnormal hemoglobin</b>	<b>327</b>
14.11 Type 3 hyperlipoproteinemia	256	REENA DAS AND PRASHANT SHARMA	
14.12 Treatment of primary lipoprotein disorders	256	18.1 Introduction	327
14.13 Management of hypercholesterolemia	256	18.2 The hemoglobin molecule: normal structure and function	327
14.14 Agents in development	260	18.3 The classification and genetics of the thalassemias	328
14.15 Management of hypertriglyceridemia	260	18.4 Gene deletions in $\beta$ thalassemia	330
References	260	18.5 Other less common, specific molecular causes of $\beta$ thalassemia	330
Further reading	265	18.6 Laboratory diagnosis of $\beta$ thalassemia	331
<b>15. Molecular medicine of diabetes mellitus</b>	<b>267</b>	18.7 Prevention of $\beta$ thalassemia	331
DHAVENDRA KUMAR		18.8 $\alpha$ Thalassemia	332
15.1 Introduction	267	18.9 Qualitative defects (structural hemoglobin variants and other abnormalities)	334
15.2 Molecular basis of glycemic homeostasis	267	18.10 Summary	337
15.3 Disorders of the glycemic regulation	270	References	337
15.4 Diagnosis of diabetes mellitus	274	Further reading	339
15.5 Overweight, obesity, and diabetes mellitus	276	<b>19. Coagulation and bleeding disorders</b>	<b>341</b>
15.6 Vitamin D and diabetes mellitus	281	INDU SINGH AND RAGHU INDER SINGH	
15.7 Inherited monogenic diabetes mellitus	283	19.1 Introduction	341
15.8 Management of diabetes mellitus	285	19.2 Genetic basis of thrombosis	343
15.9 Summary	286	19.3 Genetic basis of the bleeding disorders	345
References	287	19.4 Structural defects of the vascular system	348
Further reading	288	19.5 Inherited defects of platelets	349
<b>16. Molecular genetic management of epilepsy</b>	<b>289</b>	19.6 Conclusion	349
DAVID LEWIS-SMITH, DONALD P. CRAIG, NAOMI J.P. THOMAS, KHALID HAMANDI AND RHYS H. THOMAS		Conflict of interest	350
16.1 Introduction	289	Author's roles	350
16.2 What are seizures?	289	References	350
16.3 What is epilepsy?	290	<b>20. Molecular and genomic basis of bronchial asthma</b>	<b>353</b>
16.4 Evidence for the genetic basis of epilepsies	291	ANUBHUTI GUPTA, SAMARPANA CHAKRABORTY AND ANURAG AGRAWAL	
16.5 The genetic architecture of epilepsies	292	20.1 Introduction	353
16.6 Mitochondrial epilepsies	298	20.2 Genomics of bronchial asthma	353
16.7 Progressive myoclonic epilepsies	299	20.3 Genetic and genomic studies in bronchial asthma	354
16.8 Pharmacogenetics of epilepsy	299	20.4 Management and treatment of bronchial asthma	363
16.9 Molecular genetic testing strategies for epilepsy	300	20.5 Conclusion	363
16.10 Summary	304	References	364
References	304	<b>21. Molecular systems in inflammatory bowel disease</b>	<b>367</b>
<b>17. The human leukocyte antigen system in human disease and transplantation medicine</b>	<b>309</b>	SAAD PATHAN	
GAURAV SHARMA, AJAY KUMAR BARANWAL AND NARINDER KUMAR MEHRA		21.1 Introduction	367
17.1 Introduction	309	21.2 Complex clinical predisposition with complex complications	367
17.2 Human leukocyte antigen system	310		
17.3 Human leukocyte antigen and disease	312		
17.4 Human leukocyte antigen expression: an explanation for disease development	314		
17.5 Human leukocyte antigen and organ transplantation	317		

21.3 The identification of the <i>NOD2</i> gene	374	24. Molecular mechanisms in cancer	
21.4 <i>NOD2</i> and innate immunity	375	susceptibility—lessons from inherited cancers	423
21.5 Major histocompatibility complex (6p21)	377	SHIRLEY V. HODGSON	
21.6 The causative genome variants and functional implications	378	24.1 Introduction	423
21.7 Autophagy	379	24.2 Inherited and familial cancer	423
21.8 Adaptive immune system	379	24.3 Oncogenes and tumor suppressor genes	424
21.9 Mucosal barrier function	379	24.4 DNA repair genes	424
21.10 The parallels and paradoxes with other diseases	380	24.5 Cancer family syndromes	425
21.11 Clinical implications and translation	380	24.6 Genetic imprinting and cancer	426
21.12 Conclusion	381	24.7 Complex cancer genomics	427
References	381	24.8 Inherited susceptibility to leukemia	427
Further reading	388	24.9 Tumor markers in circulating blood	430
		24.10 Genetic counseling for inherited cancer susceptibility	430
22. Molecular biology of acute and chronic inflammation	389	24.11 Diagnostic and predictive genetic testing for cancer	431
DHAVENDRA KUMAR		References	434
		Further reading	436
22.1 Introduction	389		
22.2 Molecular pathology of acute inflammation (sepsis and trauma)	390	25. Clinical molecular nephrology—acute kidney injury and chronic kidney disease	437
22.3 Molecular pathology of chronic inflammation	391	SHIVARAM HEGDE	
22.4 Age-associated chronic inflammation	392	25.1 Introduction	437
22.5 Molecular diagnosis and treatment of chronic inflammatory diseases	394	25.2 Acute kidney injury	437
22.6 Summary	400	25.3 Chronic kidney disease	438
Acknowledgments and disclaimer	401	25.4 Clinical renal genomic medicine	440
References	401	25.5 Molecular basis of kidney transplantation	442
		25.6 Conclusion	442
23. Molecular basis of susceptibility and protection from microbial infections	403	References	443
RASHID MERCHANT, PEZAD DOCTOR AND AMI VARAIYA			
23.1 Introduction	403	26. Molecular basis of chronic neurodegeneration	447
23.2 Understanding host genetic variation of susceptibility to infectious disease	404	MEGHA AGRAWAL	
23.3 Clinical relevance of human leukocyte antigen gene variants in HBV infection	405	26.1 Introduction	447
23.4 Human leukocyte antigen gene variants and susceptibility and persistence of HBV infection	406	26.2 Neurodegenerative disease clinical case studies and molecular systems underpinning the clinical scenario	448
23.5 Immunogenetics and microbial infection	407	26.3 Molecular pathology of neurodegenerative diseases	453
23.6 Host genetic susceptibility to human papillomavirus infection and development of cervical cancer	408	26.4 Application of molecular diagnostics in neurodegeneration	456
23.7 Host genetics of Epstein–Barr infection	409	26.5 Summary	458
23.8 Dengue viral infection	410	References	459
23.9 Genetic susceptibility of humans to hantavirus infection	411		
23.10 Host factors and genetic susceptibility to intracellular bacteria	412	27. Molecular basis of movement disorders	461
23.11 Host genetic susceptibility and protection from fungal infections	415	CLAIRE MacIVER, DAFYDD MORGAN AND KATHRYN J. PEALL	
23.12 Malaria	417	27.1 Introduction	461
23.13 Genetics of susceptibility to enteral pathogens	419	27.2 Parkinson's disease	461
23.14 Conclusion	420	27.3 Dystonia	464
References	420	27.4 Ataxia	472
		27.5 Ataxia: key learning points	480
		27.6 Other movement disorders	480
		27.7 Conclusion	483
		References	483



28. Molecular pathology in neuropsychiatric disorders	487	31. Personalizing medicine with pharmacogenetics and pharmacogenomics	521
GEORGE KIROV AND ELLIOTT REES		N.J. GOGTAY, U.M. THATTE, R. RAVI AND D. BOSE	
28.1 Introduction	487	31.1 Introduction and historical perspective	521
28.2 Copy number variation in psychiatric disorders	487	31.2 The Human Genome Project	522
28.3 Cognitive function among copy number variation carriers	489	31.3 The introduction of the field of personalized medicine/precision medicine	523
28.4 Penetrance of copy number variations	489	31.4 Genetic and molecular basis of the individual drug-response variation	523
28.5 Results from genome-wide association studies	490	31.5 Tailoring or individualizing drug therapy—select examples of application of pharmacogenomics in clinical practice	526
28.6 High-throughput sequencing studies	492	31.6 Validation of genetic/genomic information for drug response	528
28.7 Pathway analysis	493	31.7 Clinical implementation of pharmacogenetics and pharmacogenomic information	528
28.8 Conclusions	493	31.8 Clinical pharmacogenetics implementation consortium—helping clinicians understand and apply pharmacogenetic information in practice	530
References	494	31.9 Pharmacogenomics and drug development—novel study designs in precision medicine	531
29. Targeted molecular therapy: the cancer paradigm	497	31.10 Ethical issues in genomic medicine	532
DYFUD MARK DAVIES		31.11 Resources to collect and curate pharmacogenetic variants	532
29.1 Introduction	497	31.12 The future of pharmacogenetics, pharmacogenomics, and personalized medicine	533
29.2 Oncogene addiction	497	References	533
29.3 Synthetic lethality	499	Further reading	534
29.4 Histology agnostic treatment	499		
29.5 Limitations of molecularly targeted therapy in cancer	500		
29.6 Making common cancer rare	501		
29.7 Conclusion	501		
References	502		
30. Gene, genome, and molecular therapeutics	505	32. Integrated genomic and molecular medicine	535
DHAVENDRA KUMAR		DHAVENDRA KUMAR	
30.1 Introduction	505	32.1 Introduction	535
30.2 Recombinant protein drugs and vaccines	505	32.2 Genetic, genomic, and molecular revolutions in medicine	535
30.3 Stem-cell therapy	509	32.3 Evidence-based, precision, and personalized medicine	537
30.4 Gene therapy	510	32.4 The stratified medicine	539
30.5 Antisense oligonucleotides	512	32.5 Integrated genomic and molecular medicine	541
30.6 Ribozymes	513	32.6 Summary	541
30.7 RNA interference	514	References	543
30.8 Aptamers	515		
30.9 Gene and genome editing	516	<b>Glossary—molecular medicine</b>	<b>545</b>
30.10 Summary	519	<b>Index</b>	<b>555</b>
Disclaimer and acknowledgments	519		
References	520		



# List of contributors

---

- Anurag Agrawal** CSIR Institute of Genomics and Integrative Biology, Delhi, India; Academy of Scientific and Innovative Research, Delhi, India; Baylor College of Medicine, Houston, Texas, United States
- Megha Agrawal** Department of Bioengineering, University of Illinois at Chicago, Chicago, IL, United States; USA Prime Biotech LLC & Biotechnology, Kiosk Florida, United States
- Saleh AlBanyan** Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada
- Ajay Kumar Baranwal** Command Hospital - Armed Forces Medical College, Pune, India
- Elijah R. Behr** Department of Cardiovascular Medicine, St. George's Hospital Medical School, London, United Kingdom
- Deepak Bhatnagar** Cardiovascular Research Group, Division of Diabetes, Endocrinology & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom
- D. Bose** Department of Clinical Pharmacology, Seth GS Medical College and KEM Hospital, Mumbai, India
- María V. Busi** Centre for Photosynthetic and Biochemical Studies (CEFOBI-CONICET), National University of Rosario, Argentina
- Samarpana Chakraborty** CSIR Institute of Genomics and Integrative Biology, Delhi, India
- David Chitayat** The Prenatal Diagnosis and Medical Genetics Program, Department of Obstetrics and Gynecology, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada; Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada
- Donald P. Craig** Department of Neurology, Royal Victoria Infirmary, Newcastle University, Newcastle-Upon-Tyne, United Kingdom
- Reena Das** Department of Hematology, Postgraduate Institute of Medical Education and Research, Chandigarh, India
- Dyfed Mark Davies** Department of Oncology, Singleton Hospital, Swansea, Wales, United Kingdom; Division of Cancer & Genetics, Institute of Medical Genetics, Cardiff University, Cardiff, United Kingdom
- Pezad Doctor** Department of Pediatrics, Nanavati Super Speciality Hospital, Mumbai, India
- Tatiana Garofalidou** William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom
- Koumudi Godbole** Clinical Geneticist, Department of Genetics, Deenanath Mangeshkar Hospital and Research Center, Pune, India
- N.J. Gogtay** Department of Clinical Pharmacology, Seth GS Medical College and KEM Hospital, Mumbai, India
- Diego F. Gomez-Casati** Centre for Photosynthetic and Biochemical Studies (CEFOBI-CONICET), National University of Rosario, Argentina
- Anubhuti Gupta** CSIR Institute of Genomics and Integrative Biology, Delhi, India
- Luciana Amaral Haddad** Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil
- Khalid Hamandi** Cardiff University Brain Research Imaging Centre, Cardiff University, Cardiff, United Kingdom; Department of Neurology, University Hospital of Wales, Cardiff, United Kingdom
- Shivaram Hegde** Children Kidney Centre, University Hospital of Wales, Cardiff, United Kingdom
- Shirley V. Hodgson** Cancer Genetics, Department of Clinical Genetics, St. George's University of London, London, United Kingdom; Cancer Genetics Consultant, Leicester Royal Infirmary, Leicester, United Kingdom

- Atul Kalhan** Diabetes & Endocrinology, Royal Glamorgan Hospital, Llantrisant, United Kingdom
- George Kirov** Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, United Kingdom
- Lotte Kleinendorst** Department of Clinical Genetics, Amsterdam University Medical Centers, Amsterdam, The Netherlands
- Dhavendra Kumar** Institute of Medical Genetics, University Hospital of Wales, Cardiff, United Kingdom; Genomic Policy Unit, University of South Wales, Pontypridd, United Kingdom; William Harvey Research Institute, Queen Mary University of London, London, United Kingdom
- David Lewis-Smith** Department of Neurology, Royal Victoria Infirmary, Newcastle University, Newcastle-Upon-Tyne, United Kingdom; Institute of Neuroscience, Newcastle University, Newcastle-Upon-Tyne, United Kingdom
- Claire MacIver** MRC Neuropsychiatric Genomics Institute, Cardiff University, Cardiff, United Kingdom
- Narinder Kumar Mehra** National Chair and Former Dean (Research), All India Institute of Medical Sciences, New Delhi, India
- Rashid Merchant** Department of Pediatrics, Nanavati Super Speciality Hospital, Mumbai, India
- Dafydd Morgan** MRC Neuropsychiatric Genomics Institute, Cardiff University, Cardiff, United Kingdom
- Patricia B. Munroe** William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom; National Institutes for Health Research, Barts Cardiovascular Biomedical Research Centre, Queen Mary University of London, London, United Kingdom
- Julian O.M. Ormerod** Department of Cardiology, Milton Keynes General Hospital, Milton Keynes, United Kingdom; Department of Cardiovascular Medicine, University of Oxford, United Kingdom
- Saad Pathan** University of Cambridge Metabolic Research Laboratories/NIHR Cambridge Biomedical Research Centre, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom
- Kathryn J. Peall** MRC Neuropsychiatric Genomics Institute, Cardiff University, Cardiff, United Kingdom
- R. Ravi** Department of Clinical Pharmacology, Seth GS Medical College and KEM Hospital, Mumbai, India
- Elliott Rees** Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, United Kingdom
- Jonathan Schofield** Cardiovascular Research Group, Division of Diabetes, Endocrinology & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom; Manchester Diabetes Centre, Manchester Royal Infirmary, Manchester University Hospitals NHS Foundation Trust, Manchester, United Kingdom
- Parin Shah** Department of Cardiology, University Hospital of Wales, Cardiff, United Kingdom
- Gaurav Sharma** Department of Translational and Regenerative Medicine, Postgraduate Institute of Medical Education and Research, Chandigarh, India
- Prashant Sharma** Department of Hematology, Postgraduate Institute of Medical Education and Research, Chandigarh, India
- Pierre Sinajon** The Prenatal Diagnosis and Medical Genetics Program, Department of Obstetrics and Gynecology, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada; Department of Pediatrics, Division of Genetics, University of Rochester Medical Center, NY, United States; Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada
- Indu Singh** School of Medical Science, Griffith University, Southport, QLD, Australia
- Raghu Inder Singh** RP WELLTAR Hospital, Bastara, Karnal, India
- Handrean Soran** Cardiovascular Research Group, Division of Diabetes, Endocrinology & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom
- Alan Graham Stuart** Honorary Associate Professor of Sports and exercise Cardiology, University of Bristol, Bristol, United Kingdom; Consultant Cardiologist, Bristol Heart Institute, Bristol, United Kingdom
- U.M. Thatte** Department of Clinical Pharmacology, Seth GS Medical College and KEM Hospital, Mumbai, India
- Naomi J.P. Thomas** Wellcome Trust Centre for Mitochondrial Research, Newcastle University, Newcastle-Upon-Tyne, United Kingdom
- Rhys H. Thomas** Department of Neurology, Royal Victoria Infirmary, Newcastle University, Newcastle-Upon-Tyne, United Kingdom; Institute of Neuroscience, Newcastle University, Newcastle-Upon-Tyne, United Kingdom; Wellcome Trust Centre for Mitochondrial Research, Newcastle University, Newcastle-Upon-Tyne, United Kingdom



- IAN Tully** Institute of Medical Genetics, Division of Cancer & Genetics, University Hospital of Wales, Cardiff, United Kingdom
- Mieke M. van Haelst** Department of Clinical Genetics, Amsterdam University Medical Centers, Amsterdam, The Netherlands
- Ami Varaiya** Department of Microbiology, Nanavati Super Speciality Hospital, Mumbai, India
- Yoshiji Yamada** Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, Tsu, Mie, Japan
- Abbas Zaidi** Department of Cardiology, University Hospital of Wales, Cardiff, United Kingdom



# About the author

---

Professor Dhavendra Kumar is an accredited clinical geneticist (United States and United Kingdom) with special interests in clinical dysmorphology (congenital malformation syndromes), clinical cardiovascular genetics (inherited conditions of heart and blood vessels), genomic and molecular medicine (precision and personalized medicine), and genomic applications in integrated health care.



He has coauthored/edited “Genetic disorders of the Indian Subcontinent,” 2004; “Genomics and Clinical Medicine,” Oxford, 2008; “Principles and Practice of Clinical Cardiovascular Genetics,” Oxford, 2010; “Oxford Specialist Handbook on Inherited Cardiac Disease,” Oxford, 2011; “Genomics and Health in the Developing World,” Oxford, 2012; “Genomic Medicine: Principles and Practice,” Oxford, 2014; “Genomics and Society,” Elsevier Cambridge (USA), 2015; “Medical Health Genomics,” Elsevier Cambridge (USA), 2016; and “Cardiovascular Genetics and Genomics: Principles and Clinical Practice,” Springer Nature, 2018. In addition, he founded and served as the Serial Editor for “Genomic and Molecular Medicine” monographs series, Morgan Claypool, 2011–18, and currently leads the multivolume series “Clinical Genomic and Precision Medicine” of Springer Nature. He is the founding Editor-in-Chief for the new genomics biomedical publications—*Genomic Medicine* (Springer-Kluwer), *Applied and Translational Genomics* (Elsevier), *Current Trends in Genomic Medicine* (Genomic Medicine Foundation, UK), and *Genomic and Molecular Cardiology* (Genomic Medicine Foundation, UK).

He founded and leads the “Genomic Medicine Foundation UK” for promoting fellowship and scholarship in genomic medicine and health care. The key initiatives of the foundation include the “Global Familial Heart Challenge (GFHC),” “Global Consortium for Genomic Education (GC4GE),” the “UK India Genomic Medicine Alliance (UKIGMA),” and the *Indo-UK Genetic Education Forum*.

**Professor Dhavendra Kumar, MBBS, MD, DCH, M.Med.Sci.,  
FRCPI, FRCP, FRCPCH, FACMG, D.Sc**

Honorary Clinical Professor, William Harvey Research Institute, Barts and London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom; Honorary Consultant in Cardiovascular Genetics, Inherited Heart Diseases Unit, Bart’s Heart Centre, St. Bartholomew Hospital, London, United Kingdom; Visiting Professor, The Genomic Policy Unit, Faculty of Life Sciences and Education, The University of South Wales, Pontypridd, Wales, United Kingdom



# Foreword

---

Shortly following the announcement of the success of the Human Genome Project (HGP) in 2003, it was suggested by several sources that the majority of common diseases would be controlled or curable within the next 20 years. However, given the complexity of most of these diseases, it soon became clear that this was rather an over-optimistic outlook.

However, a great deal of progress has been made toward an understanding of the molecular basis of disease over recent years. For example, the development of the clinical application of next-generation sequencing of both DNA and RNA has become applicable to a wide range of clinical disorders. These advances, together with a better understanding of cell biology, are already starting to play a major role in the management of human disease.

In this book, Professor Kumar, other than reviewing the fundamentals of molecular and cell biology, has described the current status of the molecular basis of a large number of common diseases. He has also considered the current status of therapeutics based at the molecular level. The latter includes new drug design, genome editing, and several other recent therapeutic aspects of molecular genetics.

Overall this is a valuable review of the current state of knowledge of the field of molecular medicine. It should be of particular value to medical students since there is now little doubt that an understanding of the molecular basis of disease will be critical for the next generation of doctors. Given the complexity of this field, this book should also be of interest to those who are organizing the teaching of medical students, particularly in the early days of their course. Should they design an introductory course on the molecular aspects of medical practice for example? Given the complexity and novelty of this new field, it will surely require rethinking regarding the early days of medical school education.

**The Late Professor Sir David Weatherall GBE, MD, FRCP, FRS**  
*Weatherall Institute of Molecular Medicine, The University of Oxford,  
Oxford, United Kingdom*



# Preface

---

Historically, biomedical sciences have gone through many phases that followed major discoveries and inventions. Inevitably, peoples, communities, and nations around that time had demanded to embrace these new advances and developments in all aspects of life, particularly in medicine and healthcare. The molecular medicine is one of the most dynamic fields of all the life sciences that encompass many disciplines of biology, biochemistry, human biology, and human genetic. The main impact of molecular medicine is evident from many advances and applications in clinical pharmacology. Undoubtedly, the knowledge of the molecular basis of human disease has led to development of novel diagnostic and therapeutic strategies to manage a number of inherited and acquired diseases. This relatively young field of molecular medicine has made a huge impact on medical education and training of general and specialist doctors and healthcare providers.



The current phase of the explosion of public interest in the human genome and related scientific developments is no exception. This is the “genome era,” and understandably, globally peoples have high hopes from eventual medical and healthcare implications, the *genomic medicine* with its many other names such as *stratified medicine*, *personalized*, and *precision medicine*. However, fundamental to all these new promising avenues of clinical medicine is the molecular basis of human disease and its many applications.

Alongside this extraordinary scientific development, the revolution in information technology and widespread use of digital communication tools have created a public that is increasingly demanding that the present-day and future generation of doctors and healthcare professionals are experts in the diagnostic, therapeutic, and preventive implications of applied genomic and molecular medicine. Arguably there is a considerable conceptual overlap and physical evidence that genomic and molecular medicine is not distinct but merely two ends of the spectrum. This is true as the genome provides the very basic components such as specific functional nucleic acid (DNA and RNA) sequences, promoting and regulating parts of the coding regions, evolutionary conserved widespread genomic polymorphisms and copy number variations, and hitherto unknown functional impact of the noncoding regions of the genome. Collectively, all cloned regions of the genome are molecules and then the gene product, the peptide chain, is in itself a molecule. However, in everyday practice of clinical medicine a number of molecular investigations are routinely carried, such as hemoglobin, clotting factors, immunoglobulins, lipoproteins, cholesterol, triglyceride, thyroid-stimulating hormone, cortisol, and many more. All these are the end points of genes and genome. It is thus safer to simply use the term clinical molecular medicine.

Molecular medicine in an interdisciplinary approach intricately linked to life science research. It allows the medical student and clinicians to understand normal and pathological cellular processes at the molecular level. There are many examples of molecular applications to the contemporary medical practice, which have fundamentally changed diagnostic and therapeutic applications. The focus has now shifted from simply symptomatic management to the molecular pathology of human diseases. This is aided by new drug discovery and development dependent upon unraveling of the molecular processes and pathophysiology specifically involving the molecular targets. However, the new drug development is a long-term painstaking and expensive effort that starts with basic and preclinical research followed by successive clinical trials and finally approval for therapeutic use by a statutory authority.

Molecular medicine is central to a number of undergraduate and postgraduate teaching programs worldwide. In addition, it is an integral part of well-established teaching courses, particularly related to many aspects of medicine. This textbook offers the entire spectrum of molecular medicine in a clinically oriented manner using relevant clinical cases and scenario. Conceptualization and planning of the book required scrutiny of the available evidence that provides insight to normal and pathological cellular processes at the molecular level.

This allowed to the selection of the topics that are covered in this book. The focus has been on the molecular perspective of the pathophysiology, improving the diagnosis and the targeted molecular therapy.

This is an edited textbook comprising many chapters. The broad and diverse range of topics included in this book required contributions from many experts on different areas of molecular medicine. However, all efforts are made to present a homogeneous textbook in a uniform and consistent style. This is a clinically oriented book with relevant scientific factual and applied information on related molecular systems. As far as possible, authors have included pertinent clinical cases with discussion on the molecular systems underpinning the pathophysiology and clinical manifestation. Most chapters, except for those covering the basic cellular and molecular biology, are arranged in a consistent format including the introduction, clinical cases, discussion with reflections on the molecular systems underpinning the clinical scenario, overview of the relevant molecular biology and pathophysiology, targeted molecular diagnosis and therapy, summary points, and relevant references with guide to further reading. Each chapter includes numerous color illustrations and summary of critical genes and molecules. References are kept to as low as possible with important cross-references to help the reader to understand complex interdependencies between the different fields of molecular medicine. Among many examples the monoclonal antibody Herceptin not only provides a comparatively new treatment option for aggressive breast cancer but is also a recombinant protein drug that is difficult to produce. Nevertheless, the successful application of this molecular evidence is one of the paradigmatic applications of pharmacogenetic testing. There are now many such applications for the treatment of cancer of lung, skin, and colon.

This textbook on clinical molecular medicine is set out in three sections—the fundamentals of cell and molecular biology, the molecular systems in human disease, and the molecular therapeutics. As far as possible, practically all chapters follow a set pattern as described in the preceding paragraph. Some of the novel diagnostic and therapeutic applications included form the basis of genomic and precision medicine that are essential components of clinical molecular medicine. Novel treatment modalities are included based on new drugs targeted at specific molecular level, stem-cell and somatic-cell therapy, RNA interference, and the CRISPR/Cas9 gene editing. The concluding section of the book includes a chapter on the ethical dimensions of molecular medicine.

It is hoped that readers and users of this textbook will enjoy the journey through the field of molecular medicine and share the enthusiasm for this fascinating subject. The success of this book would be reflected in building a firm base and developing skills in the practice of clinical molecular medicine.

**Dhavendra Kumar**



# Acknowledgement and Disclaimer

---

This book is the product of my long-standing obsession to look for core evidence underpinning the past, current, and future practice of clinical medicine. For several years the practice of medicine relied on conventional history and physical manifestations packaged with extremely restricted biological, environmental, and sociocultural variables. Scientific developments from 16th century onward aided hugely for the modern medicine. The concept of molecular basis of medicine got some recognition in the early 1970s. However, it was limited to few immunologic and hematologic proteins and related biomarkers. Nevertheless, it was a major shift and led to setting up of few leading centers of excellence in the United Kingdom, Europe, and North America. Among these, the Oxford's Weatherall Institute of Molecular Medicine occupies a special status, founded and led by Professor Sir David Weatherall, formerly the Regius Professor of Medicine in Oxford University. It is an iconic institution that echoes around the globe for promoting and setting standards for research, teaching and training clinicians and scientists on the molecular basis of clinical medicine. Over the years, new advances in genetics and genomics strengthened the scientific basis of clinical medicine.

I am humbled with gratitude to dedicate this book to my mentor and teacher, Sir David Weatherall, who despite ill health and severe aging problems continued to guide me from the last few days before passing away on December 8, 2018. I am lost for words for sharing my gratitude but firmly believe that dedicating this book to him might be one of the ways.

Planning and working on this book have been a mammoth task spanning just under 4 years. Many eminent medical teachers and academicians guided me in setting up the curriculum aimed at senior medical students, junior medical and surgical trainees, young clinicians, and senior colleagues. List is too long; however, Professor Doug Higgs, the current Director of the Oxford's Molecular Medicine Institute and my colleague Professor Angus Clarke in the Cardiff's Institute of Medical Genetics deserve special thanks with gratitude. There are many others in different places with whom I discussed and got soundings as the work on book progressed. Undoubtedly, contributions made by several authors and their colleagues, who took time out from their extremely busy work and personal schedule and produced top-quality chapters. I remain hugely indebted to each of them and acknowledge their invaluable contributions. In this context the publishing team of the Elsevier deserves special thanks, in particular, Megan Ashdown who supported me throughout with high-order diligence and perseverance to ensure eventual publication of this book. She became my friend and guide in the end, a bucketful of sincere thanks to Megan.

Any extremely challenging and demanding project, such as authoring and editing a high-order book, could not be accomplished without the support from family and friends. I am fortunate to have the continuous support of Anju, my closest friend and life partner of just over 40 years; her critical harsh persistent analysis on the book's progress helped me to fly out of dark and thick turbulent clouds on several occasions. My children, their spouses, and four lovely pretty grandchildren always reminded me a purpose to finish the book in the best possible manner. I am reminded that one day probably my grandchildren might browse through the book with their own perceptions and reflections.

I am aware of the fact that many chapters include text, illustrations, tables, and other material to facilitate and acknowledge work and academic contributions made by many other people. In keeping with the agreed legal and publishing ethical and moral principles and guidelines, members of the Elsevier publishing executive team and myself have sought permission to use, adapt, and reproduce such material. However, it is humanly possible that in few chapters, adequate permission might be missing or unclear. As the editor of this book, I take full responsibility and offer my unconditional apologies for this unintentional omission. Should such an instance

were noted then I would do my best to deal with as appropriately as would be possible in consultation with Elsevier, who holds the full copyrights of this book. Finally, I am fully responsible for all editorial mistakes, factual inaccuracies, inadequate or inappropriate reference citations, and missing permissions.

**Dhavendra Kumar**

## S E C T I O N 1

---

# Fundamentals of molecular medicine



# The human genome and molecular medicine

Dhavendra Kumar<sup>1,2,3</sup>

<sup>1</sup>Institute of Medical Genetics, University Hospital of Wales, Cardiff, United Kingdom <sup>2</sup>Genomic Policy Unit, University of South Wales, Pontypridd, United Kingdom <sup>3</sup>William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

## 1.1 Introduction

Toward the end of the last millennium, tremendous growth in the sophistication of the biological sciences was harnessed in medicine, the food industry, and related bioindustries. New discoveries and innovations in biological sciences during the five decades prior to the 21st century have centered on genetics and genomics. It took just over 50 years following the unraveling of the structure of the molecule of nucleic acids, the key unit of the biological life, for scientists to embark on sequencing of major organisms' entire genetic constitution or genome. The word *genome* includes *gene* and *ome*, implying complete knowledge of all genes and related elements in any single organism. Inevitably, this led to enthusiastic expansion of the whole science and thence to the emergence of *genomics* [1]. The suffix *-omics*, derived from the ancient Greek, refers to in-depth knowledge. Not surprisingly, genomics was followed by a plethora of related *-omics*, for example, proteomics, metabolomics, and transcriptomics [2]. Currently, we have over 30 such disciplines with the *-omics* suffix.

The ultimate goal of any scientific discipline is its translation for the benefit of all humans, crossing all possible barriers and boundaries. Major advances in medicine and health were only possible through understanding basic principles and mechanisms underlying disease processes. This was facilitated by rapid applications of physical and chemical sciences in medicine and health, for example, X-ray diagnostics, ultrasound diagnosis, microbiology diagnosis, immunohistochemical diagnosis, and finally, molecular diagnosis. Developments and advances in genetics have led to a better understanding of the principles governing heredity and the familial transmission of physical characteristics and diseases, better understanding of the pathophysiology of diseases, the development of new methods of clinical and laboratory diagnosis, and innovative approaches to making early diagnoses (e.g., prenatal diagnoses and newborn screening) and offering reproductive choices, including preimplantation genetic diagnoses. All these developments are now accepted within the broad fields of human genetics, medical genetics, clinical genetics, genetic medicine, and new emerging field of genomic medicine. Not surprisingly, the field remains wide open, encompassing the massive field of human genomics broadly focusing on medical and health genomics [3].

This chapter leads the book providing the basic factual information for grasping the concepts of heredity, genes, genetics, and genomics. It is expected that the reader will proceed to subsequent chapters better equipped with the introduction to genetic/genome sciences, genetic diseases, genetics and genomics in medicine, applications in public health, and specific issues related to society, ethics, and law [4].

## 1.2 Hereditary factors: genes, genetics, and genomics

The concept of *heredity* and *hereditary factors* dates back to several hundreds and probably even thousands of years. The popular Darwinian theory of natural selection rests on the core concept of the *transmission of hereditary*

factors [5]. For several thousand years, various descriptions and explanations have been put forward to define the physical shape and functional nature of hereditary factors. In the historical context the concept of the *gene* was introduced only recently as the most acceptable answer to explain one of the hereditary factors. However, it remains unclear when and by whom this term was first introduced. It does not matter, as the term *gene* (from the Greek *genos*, race) is now universally accepted and used in the context of understanding heredity and hereditary factor, and is probably the single-most important biological factor regulating biological life, ranging from single-cell organisms to multicellular mammals. Rapid and extraordinary scientific progress made during 19th and 20th centuries has led to the development of *genetics*, the science of heredity. This has now been transformed into the broader field of *genomics* that includes all genes with all possible heritable biologically active or inactive regulatory and evolutionary genetic elements, whether recent or extending back through several thousand years of life on our planet.

In biological terms, genes, genetics, and genomics are keys to procreation, development, growth, function, and survival. The health of any living organism is judged by its physical and functional existence. Thus genes, genetics, and genomics are central to all forms of biological health, including that of humans. Human health depends not only on its own *genetic* or *genomic* constitution, but on that of other organisms whose well-being is also essential to human health—for example, food (plants, fish, and animals), shelter (homes made of wood from trees), the environment (water, trees, and plants), protection (clothes from cotton and animal skin), and transportation (animals and vehicles made of wood from trees). From a medical perspective the science of genetics or genomics offers deep insight into and evidence for a number of human diseases, including infectious diseases resulting from either lack of protection and/or failure in controlling the spread of microbial infections or parasitic infestations. This chapter introduces the reader to some of the basic facts about genes, genetics, and genomics and discusses how these impact human health and that of the plants, crops, and animals necessary for human health and survival. This is obviously more relevant to millions of people in the developing and less-developed countries, where limited resources and lack of infrastructure limit the optimal use of the science of genetics and genomics in applications to eradicate poverty and ensure optimal health. The reader will find cross-references to separate chapters in the book containing detailed information and further discussion of each subject.

A detailed description of the basic principles of genetics and human genetic diseases is beyond the scope of this chapter. Some of these facts are explained in subsequent chapters and various other information resources on basic genetics and medical genetics. However, some basic principles and relevant information are outlined in this section to assist the reader with limited understanding of basic genetics.

### 1.2.1 Structure and organization of nucleic acids

Living organisms are divided into two large classes—the *eukaryotes* and *prokaryotes*. The cells of the eukaryotes have a complex compartmentalized internal structure, the nucleus; these include algae, fungi, plants, and animals. Prokaryotes, on the other hand, are single-celled microorganisms without any specific part harboring the genetic material or genome; examples include bacteria and other related microorganisms. The other types of living organisms are viruses, which are intracellular obligate parasites living in both eukaryotes and prokaryotes, and composed of short-dispersed nucleic acid [deoxyribonucleic acid (DNA) or ribonucleic acid (RNA)] sequences.

Genetic information is transferred from one generation to the next by small sections of the nucleic acid, DNA, which is tightly packaged into subcellular structures called *chromosomes*. Prokaryotes usually have a single circular chromosome, while most eukaryotes have more than two and in some cases up to several hundred. In humans, there are 46 chromosomes arranged in 23 pairs, with one of each pair inherited from each parent (Fig. 1.1A and B). Twenty-two pairs are called *autosomes*, and one pair is called *sex chromosomes*, designated as X and Y; females have two X chromosomes (46, XX) and males have an X and a Y (46, XY).

A chromosome consists of a tightly coiled length of DNA and the proteins (e.g., chromatin) that help define its structure and level of activity. DNA consists of two long strands of nucleotide bases wrapped round each other along a central spine made up of phosphate and sugar (Fig. 1.2). There are four bases: adenine (A), guanine (G) cytosine (C), and thymine (T). Pairing of these bases follows strict rules: A always pairs with T, and C with G. Two strands are, therefore, complementary to each other.

Genes are made up of specific lengths of DNAs that encode the information to make a protein, or RNA product. RNA differs from DNA in that the base thymine (T) is replaced by uracil (U), and the sugar is ribose. It acts as a template to take the coded information across to ribosomes for final assembly of amino acids into the protein

## Human chromosomes!

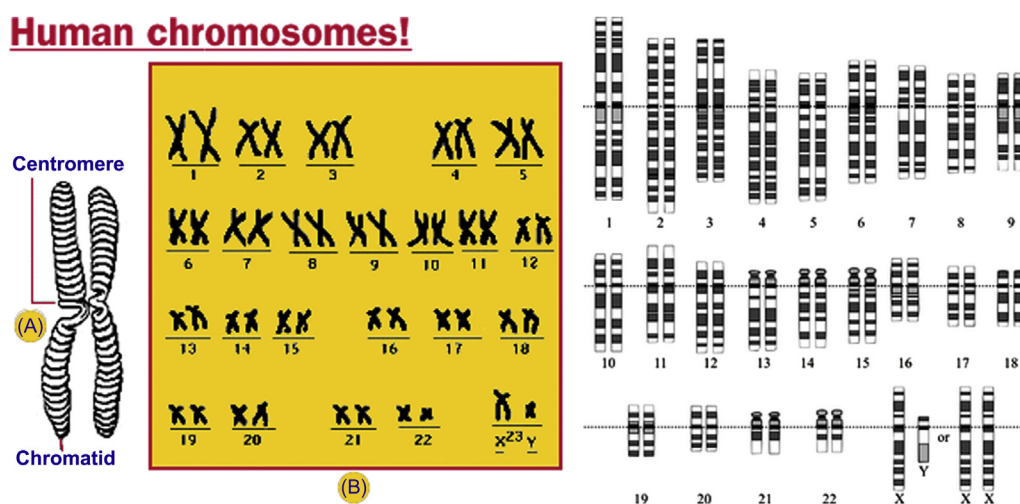


FIGURE 1.1 Human chromosomes: (A) Diploid set in a male (46, XY); (B) complete set of human chromosomes map.

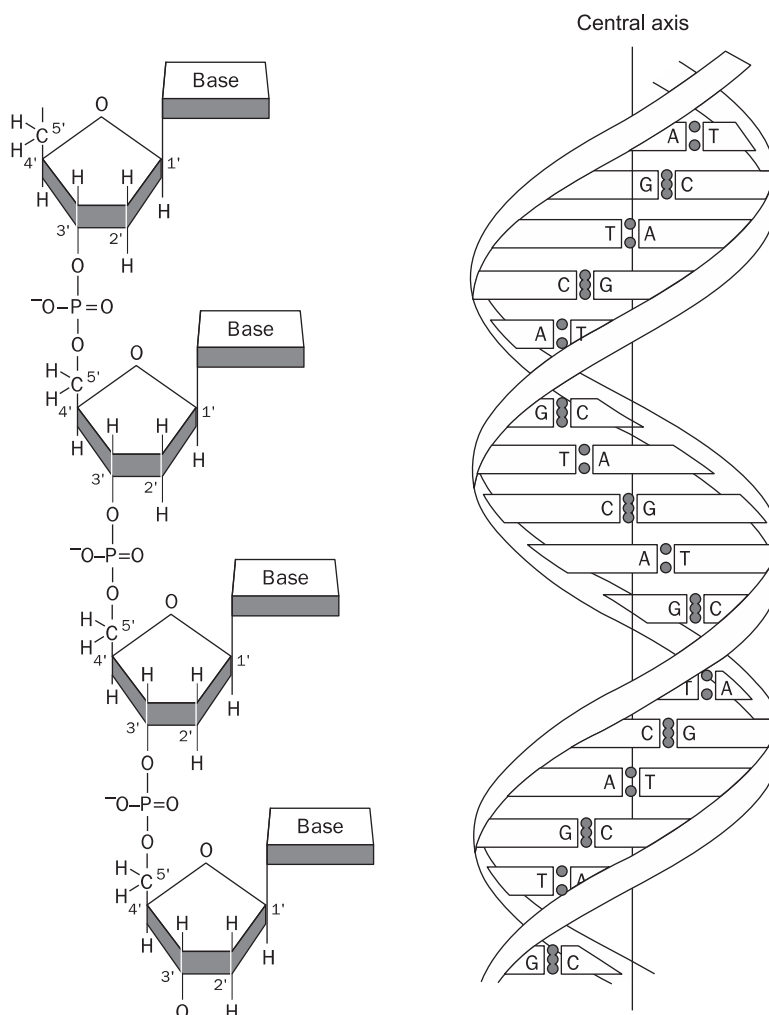
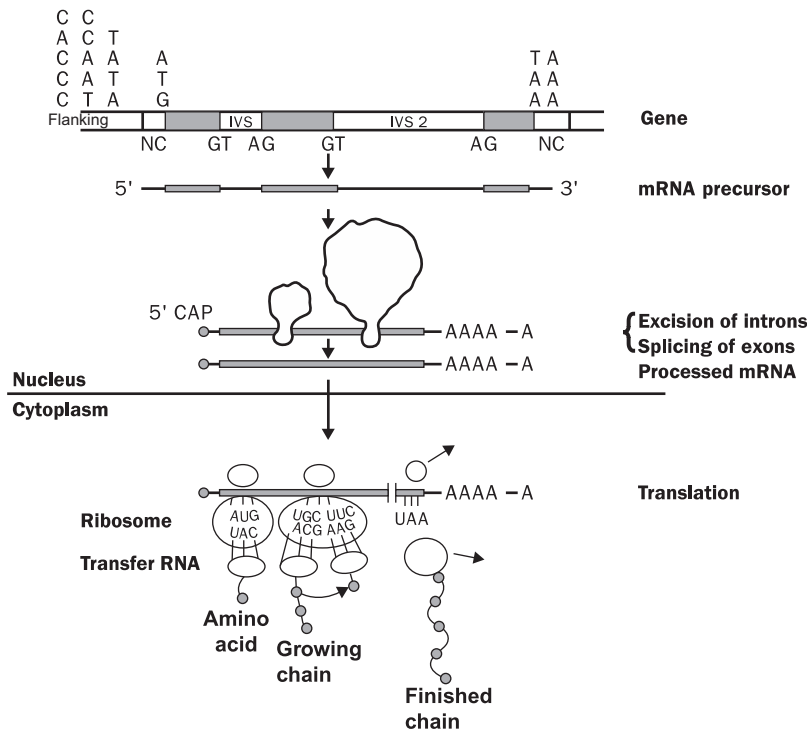


FIGURE 1.2 The Watson-Crick model of the double helix structure of the nucleic acid molecule



**FIGURE 1.3** The synthesis of the peptide chain from the coding sequences in the exon

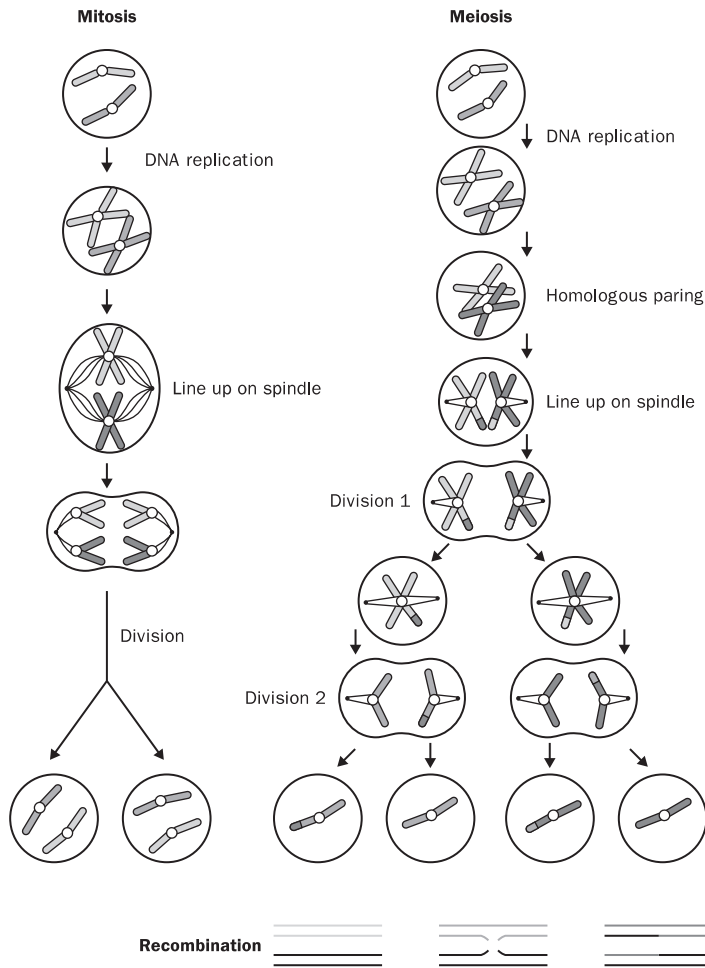
peptide chain (Fig. 1.3). The bases are arranged in sets of three, referred to as *codons*. Each codon “codes” for a specific amino acid, hence the term *genetic code*. Codons are located in *exons* that contain the coding sequences. A gene may consist of several such coding DNA segments. Exons are separated from each other by noncoding sequences of DNA, called *introns*. Although they are not yet known to be associated with any specific function, it is likely that some of these introns might be of evolutionary significance or associated with other fundamental biological functions. During the transcription of DNA, the introns are spliced out, and the exons then attach to messenger RNA (mRNA) to start the process of protein synthesis.

Proteins are one of the major constituents of the body’s chemistry. These are remarkably variable in their structure, ranging from tough collagen that forms connective tissue and bone, through the fluid hemoglobin that transports oxygen, to thousands of enzymes, hormones, and other biological effectors and their receptors that are essential for the structures and functions of the body. Each protein is made up of one or more peptide chains consisting of series of amino acids, only of which 20 occur in living organisms. The different structures and functions of proteins depend on the order of amino acids as determined by the genetic code.

DNA has the remarkable property of self-replication. The two strands of a DNA molecule separate as chromosomes divide during cell division. There are two types of cell division: *mitosis* in all body cells, and *meiosis*, which is specifically confined to the gonads in making sperm and eggs (Fig. 1.4). During mitosis, no reduction of the number of chromosomes takes place (*diploid*, or  $2n$ ), while meiosis results in half the number of chromosomes (*haploid*, or  $1n$ ). The new pairs of DNA are identical to those from which they were synthesized. However, sometimes mistakes or mutations occur. These usually result from substitution of a different base or are due to extensive structural changes to genes. In other words, any “spelling mistake” in the letters A–T or C–G could result in either absence of coded information (*nonsense mutation*) or a different message (*missense mutation*). However, not all mutations or spelling mistakes have an adverse effect (*neutral mutations*). Conversely, some changes in the genes might result in a favorable property, for example, resistance to disease or other environmental hazard. This is the basis for the gradual changes in species over millions of years of evolution. On the other hand, mutations may result in defective gene functions, leading to a disease, or susceptibility to a disease, due to qualitative or quantitative changes in the gene product, the peptide chain. However, these changes may also result from epigenetic mechanisms, abnormal RNA molecules, and posttranslational modifications (see Glossary). A brief introduction to these molecular processes is provided elsewhere in this chapter; interested readers are advised to consult dedicated texts on cell and molecular biology.

Studies on human genomic variations in different population groups and the resemblance of several genome sequences to other genomes (*comparative genomics*) have offered wide-ranging evidence to support the followers





**FIGURE 1.4** Steps in mitosis and meiosis during a eukaryotic cell division; note (*bottom*) exchange of the genetic material (recombination) through homologous pairing (Turnpenny and Ellard, 2011).

of Charles Darwin. Apart from reproduction, genes, gene-sequence variation, genomic variation, and epigenetic factors are important in growth, development, aging, and senescence. Some of these may be evolutionarily conserved across species, but relevant to human health. Mutations and alterations in several of these genomic elements are linked to a broad range of medical conditions.

### 1.3 Human genome variation and human disease

The advent of recombinant DNA technology in the 1970s revolutionized our ability to characterize and capitalize on the molecular basis of human genetic disease. This laid the foundation of eventually mapping and deciphering the DNA sequence of all the structural and functional genes of the human genome. The Human Genome Project (HGP) was, therefore, a natural progression from all previous developments in the field of human genetics. Such a mammoth task could not have been accomplished without the international collective efforts supported by generous funding from governmental and nongovernmental sources [6].

The project (HGP) has helped map and provide nucleotide sequences of around 23,000 nuclear genes, which, along with a number of other sequence variations, compose the whole human genome. Although a large number of the nuclear genes have been assigned with a structural or functional link, the precise roles of other parts of the genome are not yet fully understood. However, HGP provides the basis for “functional genomics” to explore further the genome’s functional role and understand the complex mechanisms through which genes and their products interact to affect biological function and influence disease processes. The development of new therapeutic agents is now possible on the basis of genomic arrangement and its designated functional role. This approach also helps characterize the genomes of various pathogens and other organisms, an invaluable tool in realizing the full potential of this field to improve human health [7].

### 1.3.1 Measuring genetic and genomic variation

Humans have two genomes, nuclear and mitochondrial. Normal diploid cells contain two copies of the nuclear genome and a much larger but variable number of copies of the mitochondrial genome. The nuclear genome is approximately  $2 \times 10^5$  times larger than the mitochondrial genome ( $3 \times 10^9$  vs 16,569 bp), and contains more than 1500 times the number of protein-coding genes (approximately 21,000 vs 13), including many required for mitochondrial functions. Genetic and genomic variation is abundant in both genomes. However, in general this implies to nuclear genome.

The “finished” human genome sequence was published in 2004 [7]. Table 1.1 shows the current best estimates of the size and gene content of each chromosome. These figures are for the human reference genome. There are striking differences between the nuclear and mitochondrial DNA (mtDNA) (Table 1.2). They do not correspond precisely to the genome of any actual individual, because the genomes of healthy normal individuals vary somewhat in chromosome sizes and numbers of genes, as described next. Nor is the Reference Genome in any sense an “ideal” human genome. It is simply an arbitrary and reasonably typical reference point for comparing human

**TABLE 1.1** Deoxyribonucleic acid and gene content of the reference human genome (GRCh37, February 2009).

Chromosome	Length (bp)	Protein-coding genes (known + novel)	Pseudogenes	RNA genes	Genes per Mb
1	249,250,621	2037	1131	672	8.17
2	243,199,373	1259	947	526	5.18
3	198,022,430	1066	719	430	5.38
4	191,154,276	758	698	363	3.97
5	180,915,260	874	675	343	4.83
6	171,115,067	1042	726	358	6.09
7	159,138,663	907	800	350	5.70
8	146,364,022	731	568	288	4.99
9	141,213,431	803	714	260	5.69
10	135,534,747	762	498	295	5.62
11	135,006,516	1320	774	290	9.78
12	133,851,895	1051	582	336	7.85
13	115,169,878	326	323	173	2.83
14	107,349,540	652	472	310	6.07
15	102,531,392	605	471	329	5.90
16	90,354,753	867	384	229	9.60
17	81,195,210	1197	255	273	14.74
18	78,077,248	277	56	157	3.55
19	59,128,983	1418	180	198	23.98
20	63,025,520	546	213	189	8.66
21	48,129,895	233	150	69	4.84
22	51,304,566	455	308	105	8.87
X	155,270,560	836	780	351	5.38
Y	59,373,566	53	327	44	0.89
<b>Total</b>	<b>3,286,906,385</b>	<b>21,099</b>	<b>15,520</b>	<b>11,960</b>	<b>6.42</b>

The overall totals are derived from a slightly different analysis from the individual chromosome totals, so the figures do not exactly add up. RNA, Ribonucleic acid.

Data from Ensembl Release 66, March 20, 2012.

**TABLE 1.2** Comparison of the human nuclear and mitochondrial genomes.

	Nuclear genome	Mitochondrial genome
Size (bp)	$3 \times 10^9$	16,659
Topology	23 Linear molecules	1 Circular molecule
Number of genes	Approximately 21,000	37
% Coding sequence (incl. genes for functional RNAs)	Approximately 1.4	93
Average gene density	Approximately 1 per 125 kb (variable)	1 per 0.45 kb
Introns	Average 8 per gene (variable)	None
Repetitive DNA	Approximately 50%	None

DNA, Deoxyribonucleic acid; RNA, ribonucleic acid.

genome sequences. Uncertainties in the figures relate primarily to the highly repetitive DNA of centromeres and telomeres and to the number of RNA genes, which are difficult to identify from sequence data.

The most direct way to measure genetic differences, or genetic variation, is to estimate how often two individuals differ at a specific site in their DNA sequences—that is, whether they have a different nucleotide base pair at a specific location in their DNA. First, DNA sequences are obtained from a sample of individuals. The sequences of all possible pairs of individuals are then compared to see how often each nucleotide differs. When this is done for a sample of humans, the result is that individuals differ, on average, at only about one in 1300 DNA base pairs. In other words, any two humans are about 99.9% identical in terms of their DNA sequences.

During the past several years, a new type of genetic variation has been studied extensively in humans: *copy-number variants* (CNVs)—DNA sequences of 1000 base pairs or larger are fairly distributed across the genome [8]. In some instances, CNVs could be deleted, duplicated, or inverted in some individuals with mild phenotypic effects. Several thousand CNVs have been discovered in humans, indicating that at least 4 million nucleotides of the human genome (and perhaps several times more) vary in copy-number among individuals. CNVs thus are another important class of genetic variation and contribute to at least an additional 0.1% difference, on average, between individuals. Despite significant progress, the medical and health implications of CNVs are not entirely clear [9].

Comparisons of DNA sequences can be done for pairs of individuals from the same population or for pairs of individuals from different populations. Populations can be defined in various ways: one common way is to group individuals into populations according to a continent of origin. Using this definition, individuals from different populations have roughly 10%–15% more sequence differences than do individuals from the same population [this estimate is approximately the same for both *single-nucleotide polymorphism* (SNPs)—see next—and CNVs]. In other words, people from different populations are slightly more different at the DNA level than are people from the same population. The slightness of this difference supports the conclusion that all humans are genetically quite similar to one another, irrespective of their geographic ancestry [10].

Because it is still fairly expensive to assess DNA sequences on a large scale, investigators often study genetic variations at specific sites that are known to vary among individuals. Suppose that a specific site in the DNA sequence harbors an A in some individuals' DNA sequences, and a G in others. This is a SNP, where *polymorphism* refers to a genetic site that exists in multiple forms. The proportion of individuals who have an A and the proportion with a G give the frequency of each form, or *allele*, and this frequency can be estimated for a sample of individuals from a population. If the frequencies of A in three different populations are 0.10, 0.20, and 0.50, the *genetic distance* between the first two populations is smaller than that between the third population and the first two. On the basis of this assessment the first two populations are genetically more similar than either is to the third. To get a more accurate picture of genetic differences, hundreds or thousands of SNP frequencies would be assessed to yield the average genetic difference among pairs of populations [11].

### 1.3.2 Genome variation and human disease

A number of genes have direct or indirect influence on most human diseases. Because individuals have different variants of genes, it follows that the risk of developing various diseases will also differ among individuals. Consider a simple example—Jim Fixx, a well-known runner and fitness enthusiast, died of a heart attack at the age of 52. Sir Winston Churchill, who was renowned for his abhorrence of exercise and his love of food, drink,

and tobacco, lived to the age of 90. It is plausible that genetic differences between Fixx and Churchill were responsible, at least in part, for the paradoxical difference in their life spans. (Indeed, Jim Fixx's father had a heart attack at the age of 35 and died of a second heart attack at the age of 43.)

Because genes are passed down from parents to offspring, diseases tend to "cluster" in families. For example, if an individual has had a heart attack, the risk that his or her close relatives, offspring, or siblings will have a heart attack is two to three times higher than that of the general population. Similar levels of increased risk among family members are seen for colon cancer, breast cancer, prostate cancer, type 2 diabetes mellitus (T2DM), and many other diseases. This clustering in families is partly the result of shared nongenetic factors (e.g., families tend to be similar in terms of their dietary and exercise habits) and partly the result of shared genes. As we have seen, populations differ somewhat in their genetic backgrounds. It is thus possible that genetic differences could be partly responsible for differences in disease prevalence. For many disorders caused by genetic changes in single genes, these differences are readily apparent. Cystic fibrosis, for example, is seen in about one in 2500 Europeans, but only in one in 90,000 Asians. Sickle-cell disease is much more common in individuals of African and Mediterranean descent than in others, although it is found in lower frequency in many other populations due to migration and intermarriage.

These differences in prevalence can be attributed to the evolutionary factors that influence genetic variation in general. *Mutation* is the ultimate source of all genetic variation. In some cases, such as hemochromatosis in Europeans and sickle-cell disease in Africans, the responsible mutations have arisen within the last few thousand years, helping to account for a fairly restricted distribution of the disease. *Natural selection* also plays a role in population differences in some genetic diseases. For sickle-cell disease and related diseases known as the thalassemias, *heterozygotes* (those who carry a single copy of a disease-causing mutation) are relatively resistant to the malaria parasite. Cystic fibrosis heterozygotes are resistant to typhoid fever, and hemochromatosis heterozygotes absorb iron more readily, perhaps protecting them against anemia. Also, the process of *genetic drift*, which is accentuated in small populations, can raise the frequencies of disease-causing mutation quickly just by chance (e.g., Ellis–van Creveld disease, a reduced-stature disorder, is unusually common among the Old Order Amish of Pennsylvania) [12]. In contrast to the effects of natural selection and genetic drift, which tend to promote population differences in disease prevalence, *gene flow* (the exchange of DNA among populations) tends to decrease differences among populations. With the enhanced mobility of populations worldwide, gene flow is thought to be increasing steadily.

These same factors can affect common diseases such as cancer, diabetes, hypertension, and heart disease, but the picture is more complex, because these diseases are influenced by multiple genetic and nongenetic factors. Common diseases do vary in frequency among populations: hypertension occurs more frequently in African-Americans than European-Americans, and T2DM is especially common among Hispanic and Native American populations [13]. Although genes clearly play a role in causing common diseases, it is less clear that genetic differences between populations play a significant role in causing differences in prevalence rates among populations. Consider another example: the Pima Native American population in the southwestern United States now has one of the highest known rates of type 2 diabetes in the world. About half of adult Pimas are affected. Yet this disease was virtually unknown in this population prior to World War II. Obviously, the Pimas' genes have not changed much during the past 50 or so years. Their environment, however, has changed dramatically with the adoption of a "Western" high-calorie, high-fat diet, and a decrease in physical exercise. In this case, it is almost certain that the rapid increase in type 2 diabetes prevalence has much more to do with nongenetic than genetic causes [14].

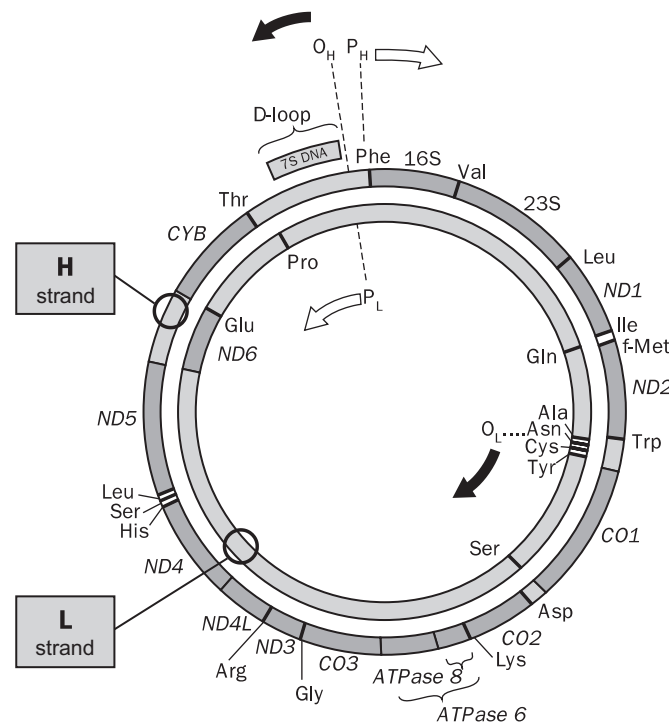
But why does a Western diet seem to have a greater effect on some populations than others? Perhaps differences in genetic background, interacting with dietary and other lifestyle changes, help account for this variation. As additional genes that influence susceptibility to common diseases are discovered, and as the roles of nongenetic factors are also taken into account, it is likely that this picture will become clearer.

---

## 1.4 The mitochondrial genome

---

The mitochondrial genome is very different from the nuclear genome (Fig. 1.5). In many respects, it has more in common with bacterial genomes than the eukaryotic nuclear genome. This is consistent with the idea that mitochondria originated as endosymbiotic bacteria within some ancestral eukaryotic cell. If this theory is correct, over the years, the mitochondria have gradually transferred more and more of their functions to the nucleus. This is evident from the fact that a number of nuclear genes encode the great majority of mitochondrial proteins. Cells contain many mitochondria (typically 100–1000; maybe 100,000 in an oocyte), so mtDNA might be formally classified among the repetitive DNA in a cell. Although the mitochondrial genome is very small compared to its nuclear counterpart, because there are many copies, mtDNA often makes up 1% or so of total cellular DNA.



**FIGURE 1.5** The mitochondrial genome—The heavy (H) and light (L) strands of the circular 16,659 bp mtDNA double helix are shown. Protein-coding genes are shaded; transfer RNA genes are shown as short lines with the name of the amino acid. There are no introns; the heavy arrows indicate the origins and directions of replication of the two strands; the light arrows show the promoters and direction of transcription of the two multicistronic transcripts that are subsequently cleaved into individual mRNAs. mtDNA, Mitochondrial DNA; RNA, ribonucleic acid. Source: Adapted with permission from Figure 9.1 of Strachan T, Read AP. *Human molecular genetics*. 3rd ed. London and New York: Garland; 2004.

As in bacteria, the mitochondrial genome is circular and closely packed with genes. There are no introns and little intergenic noncoding DNA. Some genes even overlap. A total of 24 of the 37 genes specify functional RNAs [two ribosomal RNAs (rRNAs) and 22 transfer RNA (tRNAs)]; the other 13 genes encode components of the electron transport pathway.

A short segment of the mitochondrial genome is triple-stranded of which the D-loop (displacement loop) is noncoding produced by replication forks overlapping as they travel in opposite directions around the circular DNA. The D-loop contains the only significant amount of noncoding DNA in the mitochondrial genome. Perhaps because of this, it is the location of many of the DNA polymorphisms that are such useful tools for anthropologists researching the origins of human populations. Because there is no recombination among mtDNAs, complete haplotypes of polymorphisms are transmitted through the generations, modified only by recurrent mutation. This makes mtDNA a highly informative marker of ancestry, at least along the maternal line.

mtDNA replication and transcription use nuclear-encoded polymerases. Transcription proceeds in both directions round the circle. The initial products are two large multicistronic RNAs, which are subsequently cleaved to make the individual mRNAs. All the protein components of the translation machinery are nuclear-encoded, but mitochondria exclusively encode the tRNAs, and these use a coding scheme slightly different from the otherwise universal code. There are four stop codons—UAG, UAA, AGG, and AGA: UGA encodes tryptophan, and AUA specifies isoleucine, rather than arginine as normally. Presumably, with only 13 protein-coding genes, the mitochondrial system could tolerate mutations that modified the coding scheme in a way the main genome could not.

Mutations in mtDNA are important causes of disease and perhaps also of aging [15]. Phenotypes caused by variation in mtDNA are transmitted exclusively down the maternal line (matrilineal inheritance), but most genetic diseases where there is mitochondrial dysfunction are caused by mutations in nuclear-encoded genes and so follow normal Mendelian patterns. As cells contain many copies of the mitochondrial genome, they can be heteroplasmic, containing a mix of different sequences. Unlike mosaicism for nuclear variants, heteroplasmy can be transmitted by a mother to her children.



## 1.5 Functional genomics, transcriptomics, and proteomics

Functional genomics, specifically transcriptomics, is a systematic effort to understand the function of genes and gene products by high-throughput analysis of gene transcripts in a biological system (cell, tissue, or organism) with the use of automated procedures that allow scale-up of experiments classically performed with single genes [15]. Functional genomics can be conceptually divided into *gene-driven* and *phenotype-driven* approaches. Gene-driven approaches rely on genomic information to identify, clone, and express genes, as well as to characterize them at the molecular level. Phenotype-driven approaches rely on phenotypes, either identified from random mutation screens or associated with naturally occurring gene variants, such as those responsible for mouse mutants or human diseases, to identify and clone the responsible genes without prior knowledge of the underlying molecular mechanisms [16]. The tools of functional genomics have enabled the development of systematic approaches to obtaining basic information for most genes in a genome, including when and where a gene is expressed and what phenotype results if it is mutated, as well as the identification of the gene product and the identity of other proteins with which it interacts. Functional genomics aspires to answer such questions systematically for all genes in a genome, in contrast to conventional approaches that address one gene at a time.

Within the context of functional genomics, an important part of functional arrangement of all genomes consists of areas that are external to any coding gene sequences. These play a crucial role in gene expression. Collectively this is known as epigenome. Studies in epigenetics and epigenomics have established core principles that are responsible for modulating development and differentiation and respond to external changes. Patterns of cell- and tissue-specific gene expression are established and maintained by the patterns of epigenetic marks on the genome. These consist of DNA methylation and a variety of specific covalent modifications of histones. The epigenetic marks or signatures are established by a large series of “writers”: DNA methyltransferases, histone methyltransferases and demethylases, histone acetyltransferases and deacetylases, histone kinases and phosphatases, and so on. In some cases, small RNA molecules help ensure sequence specificity. The effects on gene expression are mediated by epigenetic “readers,” including methylated DNA-binding proteins, chromodomain, and bromodomain proteins that bind methylated and acetylated histones, respectively, and a large number of other proteins [17].

As a result of epigenetic modifications, chromatin exists in a variety of epigenetic “flavors.” The basic distinction is between heterochromatin (inactive, repressed) and euchromatin (potentially active), but subtypes define transcriptional activity and regulatory elements such as promoters, enhancers, and insulators. The flavor depends on a combination of types and relative quantities of marks rather than a simple histone code.

Central to functional genomics is the complex organization of RNA molecules that occupies bulk of the intergenic parts of the genome. Apart from well-characterized coding RNA (cRNA), messenger RNA (mRNA), rRNA, and tRNA, there are many other non-cRNAs (ncRNAs) [18]. These can be divided into “classical” ncRNAs and long intergenic ncRNAs. The classical ncRNAs are small molecules, typically 16–30 nt, derived by processing much longer precursors. There has been an explosion in our knowledge of the numbers and classes of these molecules, and this is still a very active research area. The main well-established classes are as follows:

- Small nuclear RNAs form part of the spliceosomal machinery.
- Small nucleolar RNAs act as sequence-specific guides for enzymes that chemically modify specific bases in ribosomal and other RNAs.
- microRNAs control translation of many mRNAs by binding to sequences in the 3′ untranslated region.
- piRNAs (piwi-associated RNAs) act in gametes to ensure stability of the genome. There appear to be many thousands of piRNA genes, grouped in around 100 clusters.

Table 1.3 lists the numbers of genes encoding these molecules—but these are subject to major revision, because it is very difficult to identify functional ncRNAs and distinguish them from the large number of nonfunctional variants present in the genome [19].

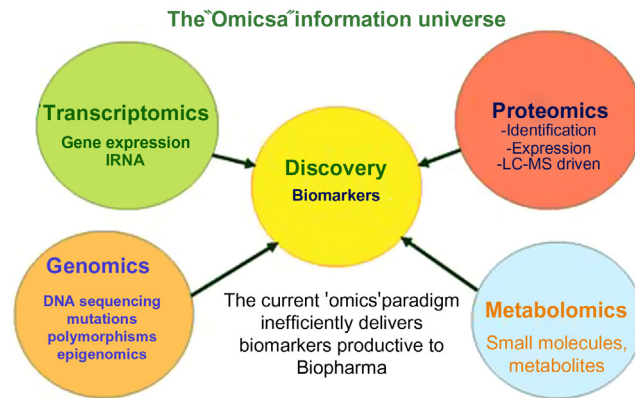
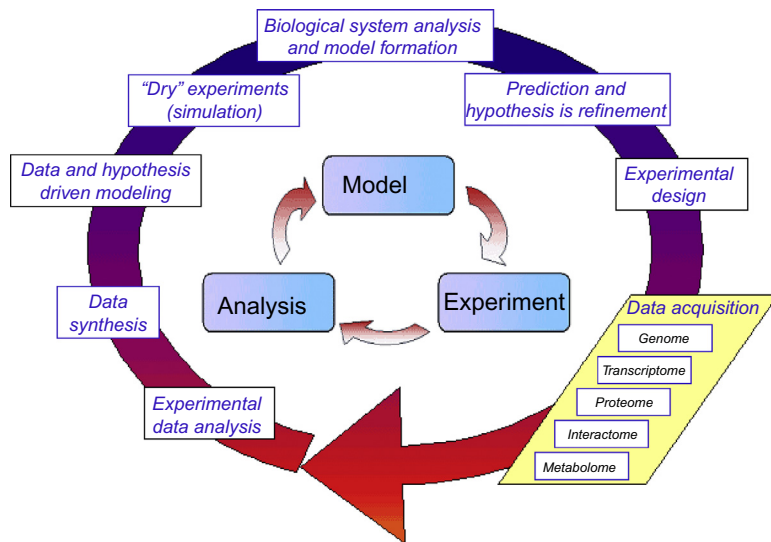
Analysis and applications of the rapid accumulation of highly sophisticated genome and proteome data necessitated development of powerful computational programs and relevant hardware tools. Storage, retrieval, and assimilation of enormous amounts of data require fast and accurate computational skills. *Bioinformatics* deals with these requirements within the broad biomedical and biotechnology sectors. There are several literature and online resources with detailed descriptions of the role and scope of bioinformatics [20].

A number of biomedical and biotechnology disciplines have emerged during the last two decades, all ending with the suffix *-omics*. *-Omics* is derived from *ome* (Greek, *omoyous*), which refers to *complete knowledge*. The ancient language Sanskrit has a similar word, *ohm*, with similar meaning and expression. A number of these

**TABLE 1.3** Ribonucleic acid (RNA) genes in the human genome.

RNA species	Number of functional genes
Ribosomal RNA	150–200
Transfer RNA	496
snRNA	91
snoRNA	375
miRNA	1733
piRNA	114 Clusters

*miRNA*, MicroRNAs; *piRNA*, piwi-associated RNAs; *snoRNA*, small nucleolar RNAs; *snRNA*, small nuclear RNAs. Griffiths-Jones [miRNAblog.com](http://miRNAblog.com), [piRNAbank.ilab.ac.in](http://piRNAbank.ilab.ac.in).

**FIGURE 1.6** The "OMICS" paradigm, showing four major branches.**FIGURE 1.7** Informatics as the central dogma for systems biology and genome sciences.

"omics" have direct or indirect links to the fundamentals of genome science and technology. A number of biological models have been developed and tested using genomic, transcriptomic, proteomic, and metabolomic approaches (Fig. 1.6). *Systems biology* refers to developing and testing biological models based on -omic sciences [21]. The central dogma of the *systems biology* is the computational analysis of complex and enormous data at all biological levels—gene, molecule, cell, tissue, organ, and whole body (Fig. 1.7).

## 1.6 Translational human genomics

The potential of applications of genome science and technology in medicine and health has led to the emergence of *genomic medicine*, a natural outcome of the tremendous progress made in medical genetics and genomics [22]. However, final endpoints in genomic medicine will largely depend upon judicious and efficacious application and utilization of the diagnostic and therapeutic potential of genome-based technologies, for example, clinical applications of microarray technology. This process requires multifaceted systematic and analytical research efforts to translate the basic scientific information into practical and pragmatic applications following the principles of good medical practice. There is no disagreement that this *translational genome research* is vital for the successful and efficient delivery of promises made by researchers and physicians behind the genomic medicine movement.

The process for translational genome research includes the participation of several researchers drawn from different disciplines. The multidisciplinary model for translational genome research is widely accepted and includes several key elements. Informatics and computational networks remain the core element for translational genomics research and systems biology [23]. A framework for the continuum of multidisciplinary translation research is recommended to utilize previous research outcomes in genomics and related areas of health and prevention [24].

The whole process of *translational genome research* includes four phases and revolves around the development of evidence-based guidelines. Phase 1 translation (T1) research seeks to move a basic genome-based discovery into a candidate health application, such as a genetic test or intervention. Phase 2 translation (T2) research assesses the value of genomic applications for health practice, leading to the development of evidence-based guidelines. Phase 3 translation (T3) research attempts to move evidence-based guidelines into health practice through delivery, dissemination, and diffusion research. Phase 4 translation (T4) research seeks to evaluate the “real world” health outcomes of a genomic application in practice. It is important to appreciate that the whole process of translation research leading to evidence-based guidelines is a dynamic one, with considerable overlap between the different stages. The process should be able to accommodate new knowledge that will inevitably arrive during translation research.

The role of translational genome research, including that of clinical trials, is crucial in developing evidence-based good-practice guidelines [25]. The aim should be to obtain vital genetic and genomic information, including laboratory material for research, from the patient, family, and community, and then use this scientific data and information for clarification and ratification. The outcomes of translational gene research should be valid and deliverable in the clinic for diagnostic and therapeutic applications [26].

## 1.7 Human genomics for socioeconomic development

During the last decade, rapid progress has been made in harnessing the huge potential of genome science and technology for its economic and health benefits globally, in particular in less- and least developed nations [27]. Apart from the World Health Organization, other international and national institutions engaged in this endeavor include the Human Genome Organization, Organization for Economic Cooperation and Development, the McLaughlin-Rotman Center for Global Health (The University of Toronto, Canada), the Mexican Health Foundation, the Beijing Genomics Institute, the Department of Science and Technology (Government of India), and many more. All these institutions are focused on supporting and exploiting the huge potential of genomic technologies and related bioinformatics developments on the global economy and on health [28]. The impact of genome sciences and technologies will manifest in the following wide-ranging areas:

- Personalized medicine and health approaches that will help people and societies shift the focus from “sick-care” to “well-care and prevention”
- Biotechnology methods to produce environmentally clean and efficient fuel and chemicals to accelerate transition from petroleum-based economies
- Genome-driven plant- and crop-growing methods for producing affordable food for less- and least developed economies
- Promoting genomic science and technology in animal breeding and livestock improvement
- Supporting genome research for new drug discovery and drug development for enhancing pharmaceutical efficacy
- Applications of genomic biotechnologies in the study and monitoring of environmental health



## 1.8 Conclusions

Developments in genetics and the subsequent sequencing of the human and other genomes have provided us with an opportunity to review the role of genes and genomes in all aspects of health and disease. Human health, including causation of disease, is not exclusively dependent on the human genes and genome. Evolutionary links with other genomes and ecologically relevant and beneficial parts of other genomes play crucial roles in the maintenance of human health and, to some extent, in morbidity and mortality. Understanding genomes of microbes, parasites, animals, plants, and crops is an acknowledged priority of current biomedical and biotechnology research.

Conventionally, the causation of human disease includes malformations, trauma, infection, immune dysfunction, metabolic abnormality, malignancy, and degenerative conditions associated with aging. Genetic factors have long been recognized in all of these disease groups. The traditional genetic categories of diseases include chromosomal disorders, single-gene or Mendelian diseases, and several forms of multifactorial/polygenic conditions. In addition, somatic genetic changes and mutations of the mitochondrial genome probably account for a small, albeit important, number of diseases. These groups of disorders are well recognized and have an established place in the classification of human disease. Recent developments in genome research have provided vast data indicating different genomic mechanisms to explain complex pathogenesis in some disorders. The spectrum of these disorders is wide and includes both acute and chronic medical and surgical diseases. Perhaps it is reasonable to identify these disorders on the basis of underlying molecular pathology, including genomic imprinting, genomic rearrangements, and gene–environment interactions involving multiple genes and genomic polymorphisms.

This chapter has reviewed the fundamental aspects of genetics and genomics that are closely related to human health and morbidity. The genomic approaches to understanding and managing human disease are rapidly being incorporated in the practice of clinical medicine, rapidly emerging as *genomic medicine*. In addition, applications of genome science and technology are also reforming biotechnologies in a number of industries, including pharmaceutical, agricultural, and ecological bioengineering. The enormous impact of genome sciences and technologies on the economy of the developing world will be judged on applications in a number of areas, including biofuels, accelerated breeding of crops and livestock, personalized health products, pharmaceutical efficacy, and genomic monitoring of environmental health.

## References

- [1] Temple LK, et al. Defining disease in the genomics era. *Science* 2001;293(5531):807–8.
- [2] Lay JO Jr, et al. Problems with the “omics”. *TrAC Trends Anal Chem* 2006;25(11):1046–56.
- [3] Feero WG, et al. Genomic medicine—an updated primer. *N Engl J Med* 2010;362(21):2001–11.
- [4] Kumar D, Weatherall DJ. *Genomics and clinical medicine*. Oxford: Oxford University Press; 2008.
- [5] Fisher RA. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press; 1999.
- [6] Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. *Science* 2003;300(5617):286–90.
- [7] Consortium IHGS. Finishing the euchromatic sequence of the human genome. *Nature* 2004;431(7011):931–45.
- [8] Conrad DF, et al. Origins and functional impact of copy number variation in the human genome. *Nature* 2009;464(7289):704–12.
- [9] Pinto D, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010;466(7304):368–72.
- [10] Freeman JL, et al. Copy number variation: new insights in genome diversity. *Genome Res* 2006;16(8):949–61.
- [11] Akey JM, et al. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res* 2002;12(12):1805–14.
- [12] Puffenberger E. Genetic heritage of the Old Order Mennonites of southeastern Pennsylvania. *Am J Med Genet, C: Semin Med Genet* 2003;121C:18–31.
- [13] Harris MI, et al. Racial and ethnic differences in glycemic control of adults with type 2 diabetes. *Diabetes Care* 1999;22(3):403–8.
- [14] Harris MI, et al. Is the risk of diabetic retinopathy greater in non-Hispanic blacks and Mexican Americans than in non-Hispanic whites with type 2 diabetes?: a US population study. *Diabetes Care* 1998;21(8):1230–5.
- [15] Yaspo M-L. Taking a functional genomics approach in molecular medicine. *Trends Mol Med* 2001;7(11):494–501.
- [16] Steinmetz LM, Davis RW. Maximizing the potential of functional genomics. *Nat Rev Genet* 2004;5(3):190–201.
- [17] Ram O, et al. Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* 2011;147(7):1628–39.
- [18] Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* 2013;20(3):300–7.
- [19] Taft RJ, et al. Non-coding RNAs: regulators of disease. *J Pathol* 2010;220(2):126–39.
- [20] Rehm B. Bioinformatic tools for DNA/protein sequence analysis, functional assignment of genes and protein classification. *Appl Microbiol Biotechnol* 2001;57(5-6):579–92.
- [21] Gehlenborg N, et al. Visualization of omics data for systems biology. *Nat Methods* 2010;7:S56–68.
- [22] Burke W, et al. Translational genomics: seeking a shared vision of benefit. *Am J Bioethics* 2008;8(3):54–6.

- [23] Karp PD, et al. Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Briefings Bioinf* 2010;11(1):40–79.
- [24] Khoury MJ, et al. Population sciences, translational research, and the opportunities and challenges for genomics to reduce the burden of cancer in the 21st century. *Cancer Epidemiol Biomarkers Prev* 2011;20(10):2105–14.
- [25] Kumar D. Clinical medicine in the genome era: an introduction. *Genomics Clin Med* 2008;53:145.
- [26] Cooper DN, et al. Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum Mutat* 2010;31(6):631–55.
- [27] Singer PA, Daar AS. Harnessing genomics and biotechnology to improve global health equity. *Science* 2001;294(5540):87–9.
- [28] Kumar D. *Genomics and health in the developing world*, vol. 62. Oxford University Press; 2012.

# Cellular structure and molecular cell biology

*Luciana Amaral Haddad*

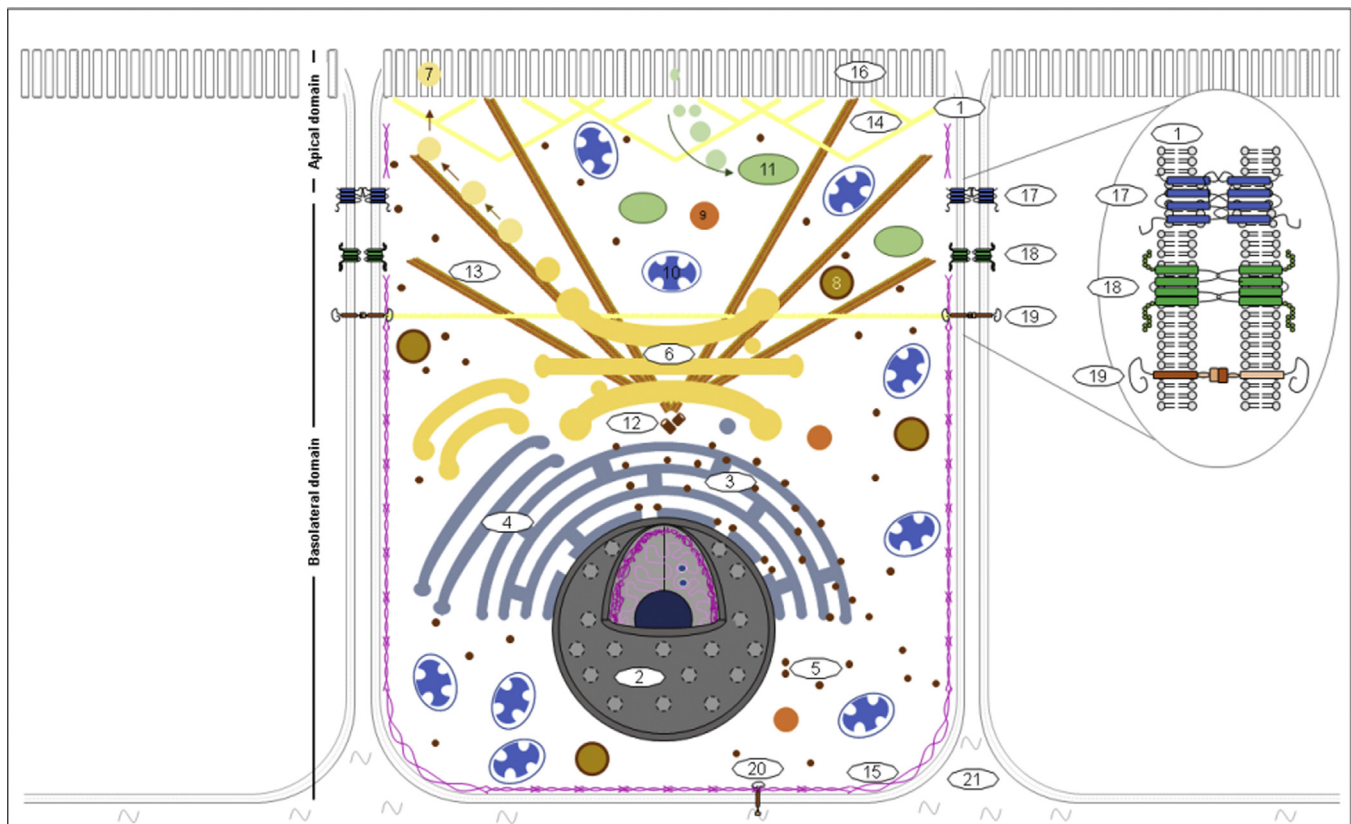
Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, São Paulo, Brazil

Eukaryotic cells present an intricate network of intracellular membranes, which defines the nucleus and other organelles with distinct biochemical composition, structure, and functions. Additional cell components, such as the cytoskeleton, ribosome, and centrosome, are unbound of membrane. Consequently, subcellular compartmentalization of membrane-bound or membrane-unbound organelles has allowed for spatial control of biological processes. The differential subcellular distribution of molecules has also permitted to evolve distinct temporal regulations of individual biological processes. In this chapter, we will introduce basic aspects to understand the cell function having the major organization of the mammalian epithelial cell organelles as a reference (Fig. 2.1). Along the text, genetic cases presenting fundamental clinical and/or laboratorial findings of a monogenic disease leading to specific organelle dysfunction will illustrate some functional roles discussed in each section.

## 2.1 Plasma membrane

The plasma membrane isolates the cytoplasm and mediates interactions between the cell and its environment. The fundamental structure of the plasma membrane and of any organelle membrane is the lipid bilayer, forming a stable barrier between two aqueous compartments. Phosphoglycerides, commonly called phospholipid, are the main constituents of lipid bilayers. Phosphoglycerides are composed of three parts: a three-carbon backbone made of one glycerol molecule; two long-chain fatty acids esterified to carbons 1 and 2 of the glycerol; and phosphoric acid esterified to glycerol carbon 3. Phosphoric acid is further esterified to a polar group such as serine, which identifies the phosphoglyceride. In the core of the bilayer, hydrophobic parts of lipid molecules from one layer face the lipophilic core of the other layer. Therefore the deepest part of the lipid bilayer consists mostly of aliphatic chains of fatty acids. As glycerol and phosphoric acid are hydrophilic, they are positioned in the external and internal faces of the plasma membrane lipid bilayer. One or more double bonds in fatty acid aliphatic chain tend to bend it contributing to the fluidity of the lipid bilayer. In each leaflet of the membrane, lateral diffusion of phosphoglycerides significantly occurs. Phospholipid bilayers are impermeable to large molecules, polar or charged compounds, or ions such as  $\text{Na}^+$  or  $\text{Cl}^-$ . On the other hand, gases and small, uncharged molecules, such as water, oxygen, carbon dioxide, urea, or glycerol, may cross the bilayer by diffusion. Therefore the lipid bilayer is a semipermeable structure.

More than 50% of the lipid content of cell membranes is phosphoglycerides (phosphatidylcholine, phosphatidylethanolamine, and phosphatidylserine). They have an asymmetric distribution between the outer and inner leaflets, with predominance of phosphatidylcholine in the outer leaflet, and preference for phosphatidylethanolamine and phosphatidylserine in the inner leaflet. The polar head of inner leaflet phosphoglycerides is negatively charged resulting in a net negative charge on the cytosolic face of the plasma membrane. The asymmetry is maintained since spontaneous flipping of lipids between leaflets is considerably low. A fifth phosphoglyceride, phosphatidylinositol, is also enriched in the inner leaflet of the plasma membrane, though in a minor proportion, and plays important roles in signaling, cell junctions and endocytosis. Besides phosphoglycerides, other lipid classes are present in the phospholipid bilayer, more specifically in the plasma membrane, such as sphingolipids (nearly 10% of lipid content) and cholesterol (30%–35%). Sphingolipids (e.g. sphingomyelin and glycolipids) are derived from sphingosine, which becomes esterified twice to a long-chain fatty acid and a polar group. The cholesterol amount associated with the phospholipid bilayer is another factor contributing to membrane



**FIGURE 2.1** Overview of epithelial cell organelles and structures. A schematic epithelial cell is presented in the center of the figure between neighboring (nondetailed) cells. Organelles and structures are indicated by numbers in the central cell, as follows: (1) plasma membrane, (2) cell nucleus, (3) rough endoplasmic reticulum, (4) smooth endoplasmic reticulum, (5) ribosomes, (6) Golgi apparatus, (7) secretory vesicle, (8) lysosome, (9) peroxisome, (10) mitochondrion, (11) endosome, and (12) centrioles. Filaments distributed in the whole cytoplasm are limited in the figure with the aim to highlight the radial distribution of microtubules (13) from microtubule-organizing centers (illustrated as centrioles of centrosomes); the branching disposition of actin microfilaments (14) in the cell cortex and their relationship with adhesive junctions; and intermediate filaments (15) interaction with hemidesmosome. Depicted specializations of the plasma membrane are microvilli (16), tight junction (17), gap junction (18), adherens junctions (19) as an example of cell-to-cell adhesive junction, and hemidesmosome (20) illustrating the cell adhesion to the extracellular matrix (21). Apical and basolateral domains of the plasma membrane are indicated.

fluidity. Lipid rafts, thicker membrane domains enriched in sphingolipids, are specialized in signaling and may be planar or slightly invaginated. The latter is known as caveolae, which are involved in endocytosis (Section 2.5).

Communication between the cell and its environment is possible because there are proteins integrally or peripherally inserted in the phospholipid bilayer. Integral proteins have a transmembrane domain with 17–25 (in average 21) nonpolar amino acids spanning across the lipid bilayer, most of the times structured as  $\alpha$ -helix. A minority of integral membrane proteins use  $\beta$ -strand to cross the lipid bilayer. Integral membrane proteins that form transporters, channels, or pumps transport small molecules, such as ions, sugar, and amino acids, across the membrane. Proteins associate peripherally with the membrane in different ways. A protein can be posttranslationally inserted in the lipid bilayer by covalent attachment to an acyl chain of the following types: isoprenoid, myristoyl, or glycosylphosphatidyl-inositol (GPI) tails. Alternatively, proteins (e.g., caveolins) may interact with lipids by electrostatic interactions or partial penetration in the bilayer. Finally, proteins may directly bind cytoplasmic domains of integral transmembrane proteins, becoming peripherally associated with the membrane. The accepted structural model for the cell membrane, the fluid mosaic, highlights the membrane fluidity due to the lipid structure and composition, and the lipid association with proteins depicting the membrane as a mosaic.

### 2.1.1 Cell signaling

One intrinsic characteristic of cells is the capability to respond to the cell environment. Different kinds of molecules transmit information between cells of multicellular organisms, acting as ligands to cell receptors. Ligands

may vary from simple gases to proteins. Although ligands such as nitric oxide or steroid hormones may diffuse through the plasma membrane and interact in the cytoplasm with respective receptors, most often ligands need to bind to the extracellular domain of its receptor, an integral membrane protein. The ligand–receptor interaction may directly associate two adjacent cells, regulating tissue differentiation or maintenance homeostasis. On the other hand, ligands may be secreted by cells residing away from the cell expressing its receptor. Endocrine signaling takes place when the ligand is a hormone secreted into the bloodstream and binds receptor in a distant cell. Paracrine signaling involves receptor-expressing cells that are in the adjacency of the cell secreting the ligand. Finally, some cells respond to signals they produced acting in autocrine signaling. In this case the same cell type expresses both the ligand and its receptor.

At least 20 distinct receptor families are known to differentially transduce the ligand signal to activate cytoplasmic relays. In spite of this diversity, signal transmission across the plasma membrane generally employs one of two different strategies. First, extracellular ligand binding may alter the conformation of the receptor, which consequently spatially modifies its cytoplasmic domain, as in guanyl cyclase receptors. Alternatively, ligand binding may trigger receptor oligomerization bringing cytoplasmic domains close enough to activate signaling cascade. One example is the tyrosine-kinase receptor dimerization activation of its own cytoplasmic kinase domain.

As exemplified, many receptors have cytoplasmic domains responsible for signal transduction. These activate downstream effectors in a binary switch form, as by phosphorylation (e.g., tyrosine-kinase or serine/threonine-kinase receptors) or guanosine 5'-triphosphate (GTP) hydrolysis (e.g., seven-helix receptors transducing through heterotrimeric GTPase proteins). Phosphorylation can activate or inhibit protein activity by directly blocking the binding site for a molecule, creating new binding sites, or changing protein conformation. As kinases act as a binary switch (on/off), phosphorylated proteins become substrate for a specific (tyrosine or serine/threonine) phosphatase, counteracting the kinase effect. Similarly, after GTP hydrolysis, exchanging guanosine 5'-diphosphate (GDP) by GTP reestablishes the active form of the GTPase (Section 2.8). A third important way to transduce ligand signals is by molecular recognition of adapter proteins, which present one or more domains acting as a platform for protein and lipid binding, some combined to residue-specific phosphorylation. Adapters generally associate with cytoplasmic domains of membrane receptors, clustering them and indirectly associating with downstream effectors.

An immediate effect of receptor–ligand transduction is the production of second messengers, small molecules that amplify the signal transduced from the primary extracellular ligand. To act as a second messenger, membrane receptor–associated signaling proteins must regulate its generation and cause an exponential increase of its amount inside the cell. Cyclic nucleotides, calcium and products from specific phosphoglyceride hydrolysis are the most common second messengers in the cell.

Different receptors coupled with heterotrimeric G proteins (GPCR) associate with adenylate cyclase on the plasma membrane. G proteins stimulate adenylate cyclase catalysis of adenosine 3',5'-cyclic monophosphate (cAMP) production from adenosine 5'-triphosphate (ATP). Binding of cAMP to the tetrameric protein kinase A (PKA) dissociates two regulatory subunits, activating the catalytic subunits. PKA-mediated protein phosphorylation controls different biological processes. In particular, PKA can translocate to the cell nucleus and phosphorylate a transcription factor (TF) (CREB), which binds to DNA-regulatory elements named cAMP-responsive element, activating transcription of regulated genes. Inactivation of cAMP by phosphodiesterase downregulates the pathway producing adenosine 5'-monophosphate (AMP). Acting in a similar way, cyclic GMP is also a relevant second messenger.

There are different lipid-derived second messengers in the cell. One example is the hydrolysis of the minor component of plasma membrane inner leaflet, phosphatidylinositol 4,5-bisphosphate (PIP<sub>2</sub>), stimulated by phospholipase C, producing two second messengers, diacylglycerol that remains in the membrane, and inositol 1,4,5-trisphosphate (IP<sub>3</sub>) released in the cytosol. Like many other lipid-derived messengers, diacylglycerol activates protein kinase C, which has different targets in the cell, and may provide positive or negative feedback to the signaling system. IP<sub>3</sub> releases calcium from intracellular store, predominantly the endoplasmic reticulum (ER). Calcium binds and activates proteins such as calmodulin, triggering downstream signaling. PIP<sub>2</sub> can also be substrate for phosphatidylinositol 3-kinase yielding phosphatidylinositol 3,4,5-trisphosphate, a second messenger that leads, among other effects, to the activation of cell growth through the mechanistic target of rapamycin (mTOR, Section 2.4).

### 2.1.2 Cell junctions

The mechanical integrity of tissues depends on the ability of cells to interact with each other and with the extracellular matrix. Cell junctions directly communicate two adjacent cells by physical contact between extracellular



domains of integral membrane proteins. Cell junctions can seal the extracellular space between two cells (tight junctions), establish selective compound transport between them (gap junctions), or make adhesive contact between cells (adhesion junctions and desmosomes) or between a cell and the extracellular matrix (hemidesmosome).

Tight junctions (*zonula occludens*) form a tight, belt-like adhesive seal in the cell circumference, selectively limiting cell migration, and diffusion of water, ions, and larger solutes in the extracellular space. In electron microscope analyses the plasma membrane of two epithelial cells appears as if they were fused together by several contacts at the tight junction level. In fact, the contact points are associations between proteins with four membrane-spanning domains, belonging to Occludin or Claudin families (Fig. 2.1). Claudin extracellular domains from adjacent cells are firmly associated forming pores with selective ion permeability (see Genetic case 1A). As Claudin family is extensive, each member has specific tissue expression and permeability patterns that classify tight junctions according to cell types. The epithelial barrier–driven paracellular transport is regulated by external stimuli, such as postprandial amino acid concentration increases in the intestine lumen.

---

*Genetic case 1A:* A term newborn from consanguineous healthy parents presents with lamellar skin desquamation and jaundice. After careful differential diagnosis among possible causes of cholestatic jaundice, the diagnosis of neonatal ichthyosis with sclerosing cholangitis syndrome was performed. This is a rare disease with autosomal recessive inheritance, caused by mutations in the *CLDN1* gene, coding for claudin-1 protein, member of a protein family essential for establishment of tight junctions.

---

Tight junctions define the border between two domains of the plasma membrane of polarized cells; the apical and basolateral domains (Fig. 2.1), which are maintained by different protein complexes, including cytoskeleton-regulating proteins. Besides the barrier function, the apical domain controls different biological functions, as secretion and absorption, as well as ionic homeostasis. For that aim the membrane apical domain develops specialized structures such as microvilli, cilia, and a system of apical transporters and channels. In addition, the cytoskeleton subjacent to the membrane apical domain, known as the cell cortex, has specific characteristics to support those membrane specializations, most remarkably, a highly regulated actin microfilament network. One example of a channel regulating ion transport across epithelial cell membrane apical domain is the cystic fibrosis transmembrane conductance regulator protein, dysfunctional in cystic fibrosis patients (see Genetic case 1B).

---

*Genetic case 1B:* At birth, a term newborn presents with meconium ileus, bilious vomiting, progressive abdominal distension, and high plasma levels of immunoreactive trypsin. Genetic testing detected a three-base pair deletion in homozygosity in the *CFTR* gene, a mutation that leads to in-frame deletion of a phenylalanine in the cystic fibrosis transmembrane conductance regulator protein, which is a chloride channel that controls ion and water secretion and absorption in epithelial tissues. *CFTR* has high expression in lungs, pancreas, and intestine. Lung function follow-up is essential.

---

The intercellular communicating gap junction allows inorganic ions and other small (less than 2000 Da) water-soluble molecules to pass directly from the cytoplasm of one cell to the neighboring cell. These intercellular channels are composed of proteins from the large Connexin family, characterized by four transmembrane domains (Fig. 2.1). A hemichannel is an assembly of six connexin protein units at the membrane. As two hemichannels keep adjacent cell plasma membranes a distance apart, a gap is established between them, having named this type of junction. Gap junctions are strikingly important in tissues that have electrically excitable cells, such as the heart, brain, and cochlea (see Genetic case 1C). Gap junctions are also essential to coordinate signaling or metabolic responses, as in the liver, skin, and gonads.

---

*Genetic case 1C:* A couple planning to start a family seeks genetic counseling as they are first-degree cousins and come from a family with high frequency of hearing impairment and consanguineous marriages. The woman's sister has profound nonsyndromic sensorineural hearing loss due to homozygous frameshifting mutation in the *GJB2* gene. She wants to know the odds of having an affected child. *GJB2* is member of a gene family encoding connexins, proteins with four membrane-spanning domains that constitute gap junctions. *GJB2* is expressed in the inner ear and is believed to be necessary for endolymph homeostasis.

---

Cells make adhesive contacts with other cells or extracellular matrix by connecting extracellular domains of adhesion molecules (Fig. 2.1). Cell adhesion molecules belong to four major protein groups: selectins, integrins, immunoglobulin superfamily, and cadherins. Among different ways the cell can adhere either to the extracellular matrix, to endothelial cells in the bloodstream, to similar cells in an epithelium or to different cell types in the same tissue; here we will consider adhesion junctions (*zonula adherens*) as example. Adhesion junctions and desmosomes are centrally composed of homotypic interactions of members of the Cadherin protein family, although heterotypic associations have also been reported. Cadherins are single-pass membrane proteins with conserved extracellular cell adhesion domains, and variable intracellular domains, which connect cadherins to actin and intermediate filaments, and adapter proteins. Classic cadherin cytoplasmic tail interacts with the adapter protein  $\beta$ -catenin, which is also a strong interactor for adenomatous polyposis coli (APC) protein, encoded by the *APC* gene. Most  $\beta$ -catenins are associated with cadherin, but free  $\beta$ -catenins are in equilibrium between binding sites in cytoplasmic APC or proteins in the cell nucleus, where it is a coactivator of gene transcription. As APC is part of a complex that directs  $\beta$ -catenin for degradation, and the *APC* gene is frequently mutated in colorectal cancer and other neoplasia, its loss of function partially relates to hyperactivation of  $\beta$ -catenin (see Genetic case 1D).

---

*Genetic case 1D:* An asymptomatic 16-year-old boy was referred for evaluation in the surgery clinics due to a palpable mass detected in his abdomen left flank. CT scans revealed two adjacent masses in the abdominal wall. Histological studies after surgical resection classified them as desmoid tumors. Additional workup was undertaken. Endoscopic examination revealed two gastric and numerous, small colonic polyps. MRI analyses disclosed bone fibrous dysplasia in the skull and maxillae. The diagnosis of familial adenomatous polyposis (FAP) was performed, and a heterozygous mutation was identified in the *APC* gene in leukocytes. Individuals with FAP have a high risk of developing colorectal cancer at an early age, and an appropriate follow-up is strongly recommended.

---

## 2.2 Cytoskeleton

Cell shape, mechanical support and locomotion, muscle contraction, and intracellular movements are biological functions essentially dependent on the cytoskeleton. Cytoskeleton provides the fundamental structure of the cell, a dynamic framework composed of three sorts of filaments and associated proteins: actin microfilaments, intermediate filaments and microtubules. Each filament is assembled *de novo* from monomeric/dimeric proteins, with additional units binding to preexisting filaments. Actin microfilaments and microtubules specifically undergo dynamic assembly and disassembly.

Among cytoskeleton filaments, actin microfilaments have the lowest diameter (7 nm) and lengthen up to few micrometers. Globular monomeric actin molecules (G-actin) form dimers or trimers based on expenditure of associated ATP. Other actin monomers are added to dimers/trimers growing into a flexible filament (F-actin). As assembly is in a head-to-tail fashion, actin microfilaments are polar, displaying barbed or pointed ends. Filament polymerization and depolymerization take place in both ends. Actin filaments may associate to other microfilaments forming bundles or networks. The cell cortex underneath the plasma membrane is a network of actin and associated proteins, which changes according to intracellular signaling. Cell contact points with other cells or the extracellular matrix in adherent junctions are sites of cytoskeleton attachment to the membrane, where actin microfilaments associate as discrete spots (focal adhesions) or accommodate large bundles of actin (stress fibers). In migratory cells the leading edge can have a broad flat shape (lamellipodium) due to extensive actin branching. In addition, cells may temporarily extend isolated thin projections supported by few actin bundles, such as in filopodia. Moreover, actin microfilaments may protrude the plasma membrane forming several finger-like projections, named microvilli, particularly enriched on the surface of cells involved in absorption, containing closely packed parallel bundles of 20–30 microfilaments. The abundance of microvilli on the apical domain of intestine epithelial cells (Fig. 2.1) defines a layer named brush border. The stereocilium, a single modified microvillus in auditory hair cells, detects sound wave vibrations.

A large and heterogeneous group of actin-associated proteins mediates microfilament extension, contraction, and formation of filament bundles or network. Some are motor proteins, such as myosin chains. Other proteins are present in filament barbed or pointed ends, in actin filament cross-linking, branching (Fig. 2.1), and attachment to membrane or to other cytoskeleton filaments. One example of actin-associated protein is dystrophin, deficient in Duchenne or Becker muscular dystrophies (see Genetic case 2A).

*Genetic case 2A:* A 7-year-old boy with apparently enlarged calf muscles was referred to a pediatric neurologist for muscle weakness, difficulty in walking, climbing stairs, and rising from sitting position. His parents were not consanguineous, and other family members were healthy. On examination, he had proximal weakness of pelvic girdle muscles, and laboratory tests disclosed 40-fold increase in serum creatine kinase and aldolase levels, strongly suggesting myopathy. Considering Duchenne muscular dystrophy, a muscle biopsy was requested.

Intermediate filaments provide and maintain mechanical support for cells individually and within a tissue through their connections with the plasma membrane, other cytoskeleton filaments, and adherent junctions (Fig. 2.1). Their diameter is about 10 nm, and they are homo- or heteropolymers of proteins from a large family displaying a central rod domain, head and tail domains, with a tissue-specific but nonexclusive expression pattern. For instance, keratins are in epithelial cells, desmin is intermediate filament normally found in muscle cells, neurofilaments (light, medium, or heavy chains) in mature neurons, glial fibrillary acidic protein in glial cells, and so on. They spontaneously assemble as parallel dimers by coiling their central rod domains. Dimers form tetramers in a shifted antiparallel manner. Tetramer overhang ends then associates in protofibers in a head-to-tail way. The parallel association of protofibers defines the intermediate filament, which adopts rope-like structures. Intermediate filaments are connected to desmosomes and hemidesmosomes by plakin adapter proteins. Loss-of-function mutations in the *DSP* gene coding for desmoplakin cause skin fragility and heart arrhythmia (see Genetic case 2B). Similar phenotypes are observed in patients with mutations in keratin genes.

*Genetic case 2B:* A preterm female newborn from third-degree relative, healthy parents presents reddish discoloration of limbs with skin peeling noticed at birth. During hospital stay, the baby developed blisters over limbs and trunk with spontaneous rupture. Similar findings were reported in two cousins. Systemic evaluation was otherwise normal, and an initial diagnosis of epidermolysis bullosa simplex is clinically suggested. Inherited epidermolysis bullosa is a group of genetically and clinically heterogeneous skin disorders with autosomal recessive inheritance. Follow-up is necessary to rule out associated nondermal clinical presentations.

Microtubules are rigid hollow rods with 25-nm diameter, fundamental for cell shape, intracellular transport of macromolecular complexes and organelles, cell migration coordination, mitosis, meiosis, and flagellum and cilium structures (Section 2.7). Microtubules are initially assembled as  $\alpha$ - and  $\beta$ -tubulin heterodimers, the protofilament, upon hydrolysis of associated GTP. Thirteen protofilaments are arrayed as a ring in microtubule-organizing centers (MTOC) (Section 2.7), such as the centrosomes, from where they will polymerize adding extra  $\alpha$ - and  $\beta$ -tubulin dimers. Microtubule ends emanating from centrosomes are plus ends, whereas those anchored at centrosomes are minus ends. Centrosomes have a close relationship to the cell nucleus. Therefore microtubules departing from centrosomes display a radial distribution in the cytoplasm (Fig. 2.1), contributing to cell shape and polarization, and proper nucleus positioning during cell migration. Microtubules can detach from centrosomes and attach somewhere else in the cell, such as the apical domain of plasma membrane in epithelial cells or axonal membrane in neurons. Microtubule plus ends are predominantly remodeled over minus ends. In neuronal dendrites, however, a significant part of microtubule minus ends is oriented away from the soma and undergoes dynamic remodeling upon synaptic function.

Microtubule-associated proteins (MAP) include kinesin and dynein motor protein families that are crucial for intracellular organelle and macromolecule transport. MAPs are also regulators of microtubule assembly and disassembly, filament bundling and cross-linking, cytoskeleton filament integration, and filament plus end association with membrane specializations. Lissencephaly, a rare neurological disorder of lack of cerebral cortex sulci, may affect boys with mutation in the X-linked gene *DCX*, which encodes a microtubule-binding protein necessary for proper microtubule function in the developing embryonic cerebral cortex (see Genetic case 2C).

*Genetic case 2C:* A 15-hour-old male baby born to healthy parents, evaluated at neonatal intensive care unit, has seizures since 3 hours of life. On physical examination, he presents receding forehead, flat occiput, hypotonia, poor cry, and poor suckling reflex. Brain MRI revealed bilaterally absent gyri and thick cortex, compatible with lissencephaly.



## 2.3 Cell nucleus and gene expression

In eukaryotes the genetic material is isolated from the cytoplasm by the nuclear membrane (Fig. 2.1), allowing for spatial organization of different gene expression steps. It is now recognized that distinct structures exist in the cell nucleus, directly related to gene transcription, pre-mRNA processing, ribosome biogenesis and transport, and mRNA transport. We present some of these structures linked to different nuclear processes of gene expression. In this and the following sections, we will discuss distinct levels of gene expression regulation from gene transcription to protein half-life control (Box 2.1).

### 2.3.1 Chromosome territories, gene transcription and the nuclear lamina

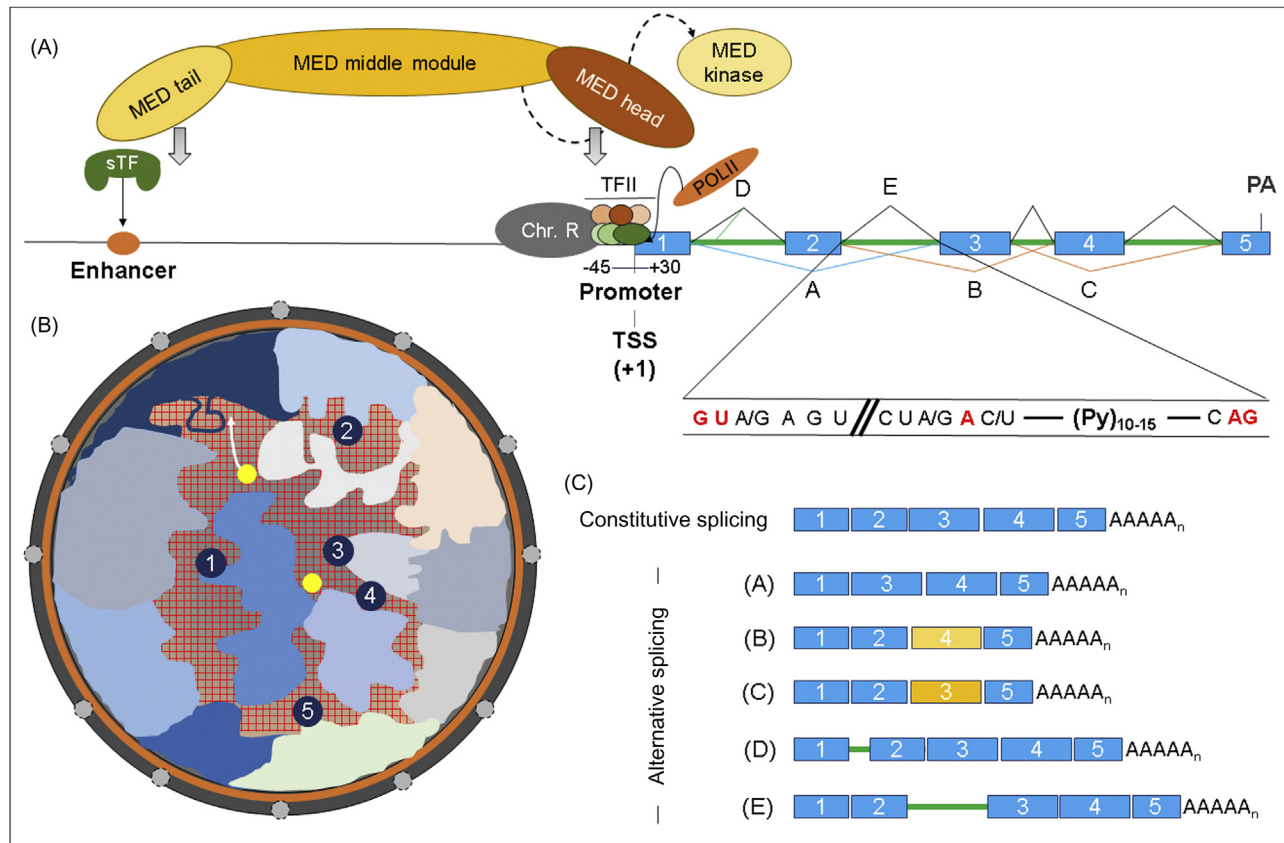
Eukaryotic genomes are spatially organized in the cell nucleus in a nonrandom three-dimensional distribution of chromosomes. It means that specific chromosomes have preferential nuclear positioning either peripherally or centrally located (Fig. 2.2). This so-called radial positioning of chromosomes in the nucleus varies between cell types from the same species. For instance, chromosome 19 may be differentially located in the nucleus of liver and kidney cells. Some intranuclear nonrandom chromosome territories are known to correlate to transcription activity. It is well established that the nuclear periphery is enriched in condensed heterochromatin, in general associated with transcription repression. The nuclear lamina is a meshwork of lamins A, B, and C intermediate filament proteins, subjacent to the nuclear envelope. In the human genome, there are more than 1000 lamina-associated chromosome domains, characterized in average with 0.5 Mb in size, poor in genes, and the contained

#### BOX 2.1

##### Eukaryote gene expression can be regulated at different levels

From chromatin through protein half-life, distinct mechanisms have evolved targeting the regulation of protein-coding genes and their expression products (mRNA and protein), as exemplified in the following:

1. *Transcriptional control of gene expression*: (a) epigenetic modifications (histone posttranslational modification such as methylation or acetylation; and gene promoter chemical modification, as by methylation); (b) alternative promoter; (c) expression of tissue-specific transcription activator or repressor that respectively binds enhancers or silencers, which are DNA-regulatory elements that may be located away from the promoter.
2. *Cotranscriptional regulation of pre-mRNA processing*: (a) alternative splicing leading to exon skipping, choice of mutually exclusive exons, intron retention, and alternative usage of 5' or 3' splicing sites; (b) alternative polyadenylation sites that may relate to alternative 3' exons; (c) lack of polyadenylation of short-living mRNA (e.g., histone mRNA).
3. *Control of mRNA cytoplasmic localization*: (a) nuclear mRNA decay by exosome; regulation of mRNA transport (b) from the nucleus to the cytoplasm, or (c) in the cytoplasm, where mRNA local concentration will directly relate to its protein product concentration (e.g., mRNA from maternal effect genes, mRNA for synaptic proteins); (d) nonsense-mediated mRNA decay, a surveillance mechanism frequently based on a pioneer translational round that signals for degradation some mRNA harboring premature translation termination codons, entering the cytoplasm.
4. *Regulation of mRNA translation initiation and half-life*: (a) general protein synthesis control by nutrient availability affecting mRNA configuration and 43S ribosome small subunit complex assembly; (b) proteins that bind to hairpins or short responsive elements mostly in untranslated regions of mRNA, specifically activating or repressing mRNA translation or stability; (c) short noncoding RNAs (miRNA) hybridized to cytoplasmic mRNA as part of the RNA-induced silencing complex leading to translational repression or mRNA degradation.
5. *Translation elongation pauses and protein folding control*: ribosome pausing during translation elongation due to cell stress, repression, rare codon occurrence, repetitive sequences, or other causes may have impact on cotranslational protein folding.
6. *Protein stability and activity*: proteasome-mediated protein degradation upon poly-ubiquitination; protein activity regulation by posttranslational modification (e.g., phosphorylation, methylation).



**FIGURE 2.2** The cell nucleus and gene expression. (A) A diagram of a hypothetical mammalian gene is presented with five exons (blue boxes) and four introns (green thick lines). A zoom in the second intron shows splicing canonical motifs in red in abbreviated conserved sequence of intron ends, and a pyrimidine-rich tract - (Py)<sub>10-15</sub> - near the 3' end. Black lines above introns indicate intron exclusion by constitutive splicing. Colored lines above and below introns indicate alternative splicing forms, corresponding to mRNA output as in (C): exon 2 skipping (A), mutually exclusive exons (B and C), alternative 5' splice site (D), and intron retention (E). The promoter (-40 to +35) is indicated as well as TSS (+1), an enhancer, and the polyadenylation signal (PA). General TFII association with promoter allows binding of RNA POLII. Chromatin remodeling complex (Chr. R) is shown upstream. A tissue sTF is able to specifically bind the enhancer. The mediator complex is responsible to integrate those protein factors. Mediator modules (head, middle, tail, and kinase) are presented. (B) Diagrammatic scheme of a eukaryotic cell nucleus showing the envelope, pores and lamina (orange), as well as chromosome territories in different colors, numbered nucleoli in early interphase, two distinct Cajal bodies (yellow), speckles (checkered), and an arrow indicating shift of snRNP to a perichromatin fibril during DNA transcription. POLII, Polymerase II; sTF, specific transcription factor; TFII, transcription factor II; TSS, transcription start site.

genes have very low or no transcription activity. Evidences suggest that cells deficient in lamin-A/C or B present misregulation of gene transcription, as in progeria, certain cardiomyopathies, lipodystrophies, and muscular dystrophies, clinical conditions grouped as laminopathies (see Genetic case 3A and Table 2.1).

*Genetic case 3A:* A young adult male, age 23, was evaluated at the cardiology clinics presenting with severe heart conduction disorder, limb muscle wasting, mild elbow tendon contractures with limited extension of both forearms, limited neck flexion, and twofold elevated creatine kinase. Emery–Dreifuss muscular dystrophy was confirmed after genetic identification of *LMNA* mutation. Lack of detection of the mutation in either parent suggested it is a new mutation in the family.

Gene transcription is controlled by a combination of factors. Chromatin basic unit is the nucleosome, a histone octamer wrapped by nearly 150 base pairs of DNA double helix. The nucleosome core may be chemically modified. Histone code represents the acetylation, mono- or tri-methylation of specific lysine residues of different

**TABLE 2.1** Examples of genes and genetic diseases associated with specific organelle dysfunction.

Disease	Gene acronym
<b>I. Cytoskeletonopathies</b>	
<b>I.a. Laminopathies (LMNA gene mutations)</b>	
Atypical progeria syndromes	LMNA
Charcot–Marie–Tooth disease, type 2B1	LMNA
Dilated cardiomyopathy, type 1A	LMNA
Emery–Dreifuss muscular dystrophy, types 2 and 3	LMNA
Familial partial lipodystrophy, Dunnigan type	LMNA
Generalized lipodystrophy	LMNA
Hutchinson–Gilford progeria syndrome	LMNA
Limb girdle muscular dystrophy, type 1B	LMNA
Mandibuloacral dysplasia	LMNA
Congenital muscular dystrophy	LMNA
Restrictive dermopathy	LMNA
<b>I.b. Associated with cell membranopathies</b>	
<i>I.b.i. Skin fragility and cardiopathy disorders</i>	
Dilated cardiomyopathy with woolly hair, keratoderma, and tooth agenesis (autosomal dominant inheritance)	DSP
Epidermolysis bullosa, lethal acantholytic (autosomal recessive inheritance) <sup>a</sup>	DSP
Naxos disease	JUP
Ectodermal dysplasia	PKP1
<i>I.b.ii. Muscle disorders</i>	
Dilated cardiomyopathy, type 1A	LMNA
Dilated cardiomyopathy, type 1D	TNNT2
Dilated cardiomyopathy, type 1R	ACTC1
Dilated cardiomyopathy, type 1Z	TNNC1
Dilated cardiomyopathy, type 1H	TPM1
Dilated cardiomyopathy, type 2A	TNNI3
Distal myopathy, type 1 (Laing distal myopathy)	MYH7
Distal myopathy, type 4	FLNC
Duchenne/Becker muscular dystrophy	DMD
Emery–Dreifuss muscular dystrophy, X-linked form	EMD
Limb girdle muscular dystrophy, type 1A	MYOT
Limb girdle muscular dystrophy, type 1B	LMNA
Limb girdle muscular dystrophy, type 1C	CAV3
Limb girdle muscular dystrophy, type 2B	DYSF
Limb girdle muscular dystrophy, type 2C	SGCG
Limb girdle muscular dystrophy, type 2D	SGCA
Limb girdle muscular dystrophy, type 2E	SGCB

(Continued)

TABLE 2.1 (Continued)

Disease	Gene acronym
Limb girdle muscular dystrophy, type 2F	SGCD
Limb girdle muscular dystrophy, type 2G	TCAP
Myofibrillar myopathy, type 1	DES
Miyoshi muscular dystrophy 1	DYSF
Nemaline myopathy, type 1	TPM3
Nemaline myopathy, type 2	NEM
Nemaline myopathy, type 3	ACTA1
Nemaline myopathy, type 4	TPM2
Nemaline myopathy, type 5	TNNT1
Nemaline myopathy, type 7	CFL2
<i>I.b.iii. Hemolytic anemia</i>	
Spherocytosis, type 1	ANK1
Spherocytosis, type 2	SPTB
Spherocytosis, type 3	SPTA1
Spherocytosis, type 5	EPB42
Elliptocytosis, type 1	EPB41
Elliptocytosis, type 2	SPTA1
Elliptocytosis, type 3	SPTB
<b>II. Channelopathies (specific channel dysfunctions)</b>	
Cystic fibrosis	CFTR
Epilepsy disorders	CACNA1H, CACNA2D2, CACNB4, CACNG2, SCN1A, SCN1B, SCN2A, SCN8A
Varying forms of heart arrhythmia	SCN5A, CACNA1C, CACNA2D1, CACNB2B
Genetic forms of migraine	ATP1A2, CACNA1A, SCN1A
<b>III. Ciliopathies (primary cilium dysfunction)</b>	
Bardet–Biedl syndrome	ARL6, BBIP1, BBS1, BBS2, BBS4, BBS5, BBS7, BBS9, BBS10, BBS12, C8ORF37, CEP290, IFT27, IFT172, LZTFL1, MKKS, MKS1, NPKP1, TTC8
Birt–Hogg–Dubé syndrome	FLCN
Joubert syndrome	AHI1, B9D1, B9D2, C5ORF42, CC2D2A, CEP41, CEP104, CEP164, CEP290, CSPP1, HYLS1, INPP5E, KIAA0556, KIAA0586, KIF7, MKS1, NPHP1, OFD1, PDE6D, PIBF1, POC1B, TCTN1, TCTN2, TCTN3, TMEM107, TMEM138, TMEM216, TMEM231, TMEM237, TMEM67, ZNF423
Merckel syndrome	B9D1, B9D2, CC2D2A, CEP164, CEP290, CSPP1, MKS1, NPHP3, RPRIP1L, TCTN2, TMEM7, TMEM216, TMEM231, TMEM237, TMEM67
Nephronophthisis	ANKS6, CEP83, DCDC2, GLIS2, IFT172, IFT81, INVS, NEK8, NPHP1, NPHP3, SDCCAG8, TCTN2, TMEM67, TTC21B, WDR19, XPNPEP3, ZNF423
Oculocerebrorenal syndrome of Lowe	OCRL
Pallister–Hall syndrome	GLI3

(Continued)

TABLE 2.1 (Continued)

Disease	Gene acronym
Polycystic kidney disease	<i>PKD1, PKD2, PKHD1</i>
<b>IV. Ribosomopathies (ribosome biogenesis dysfunction)</b>	
Bowen–Conradi syndrome	<i>EMG1</i>
Cartilage hair hypoplasia	<i>RMRP</i>
Congenital dyskeratosis, X-linked	<i>DKC1</i>
Diamond–Blackfan anemia I	<i>RPS19, RPS26, RPS27, RPS28, RPL27, TSR2, GATA1</i>
Postaxial acrofacial dysostosis	<i>DHODH</i>
Roberts syndrome	<i>ESCO2</i>
Shwachman–Diamond syndrome	<i>SBDS</i>
Treacher Collins syndrome	<i>TCOF1</i>
<b>V. Mitochondriopathies (including proteinopathies)<sup>b</sup></b>	
<b>V.a. Repeat expansion disorders</b>	
Amyotrophic lateral sclerosis/frontotemporal dementia	<i>C9ORF72</i>
Dentatorubro pallidoluysian atrophy (Haw–River syndrome)	<i>ATN1</i>
Fragile X-associated tremor and ataxia syndrome	<i>FMR1</i>
Friedreich Ataxia	<i>FXN</i>
Huntington Disease	<i>HD</i>
Myotonic dystrophy, type 1	<i>DMPK</i>
Myotonic dystrophy, type 2	<i>CNPB</i>
Oculopharyngeal muscular dystrophy	<i>PABP2</i>
Spinocerebellar ataxia type 1	<i>ATXN1</i>
Spinocerebellar ataxia type 2	<i>ATXN2</i>
Spinocerebellar ataxia type 3 (Machado–Joseph disease)	<i>ATXN3</i>
Spinocerebellar ataxia type 4	<i>PLEKHG4</i>
Spinocerebellar ataxia type 6	<i>CACNL1A4</i>
Spinocerebellar ataxia type 7	<i>ATXN7</i>
Spinocerebellar ataxia type 8	<i>ATXN8</i>
Spinocerebellar ataxia type 9	<i>SCA1</i>
Spinocerebellar ataxia type 10	<i>ATXN10</i>
Spinocerebellar ataxia type 12	<i>PP2R2B</i>
Spinocerebellar ataxia type 17	<i>TBP</i>
Spinocerebellar ataxia type 31	<i>BEAN1</i>
Spinocerebellar ataxia type 36	<i>NOP56</i>
<b>V.b. ALS and frontotemporal dementia</b>	
Most commonly mutated genes	<i>C9ORF72, FUS, SOD1 (ALS only), TARDBP, VAPB</i>
Other genes that have been suggested for the etiology of ALS/FTD due to <i>de novo</i> pathogenic variants reported	<i>ANXA11, C21orf2, CCNF, CHCHD10, CHMP2B, FIG4, HNRNPA1, MATR3, NEK1, PFN1, SIGMAR1, SQSTM1, TBK1, TIA1, TUBA4A, UBQLN2</i>
<b>V.c. Early-onset familial Alzheimer disease</b>	<i>APP, PSEN1, PSEN2</i>
<b>V.d. Familial Parkinson disease</b>	<i>ATP13A2, PARK2, PARK7, PINK1, SNCA</i>

(Continued)

TABLE 2.1 (Continued)

Disease	Gene acronym
<b>VI. Peroxisomopathies (peroxisome biogenesis disorders)</b>	
Zellweger spectrum disorder (with increasing severity: infantile Refsum disease, neonatal adrenoleukodystrophy, and Zellweger syndrome)	<i>PEX1, PEX2, PEX3, PEX5, PEX6, PEX10, PEX11B, PEX12, PEX13, PEX14, PEX16, PEX19, PEX26</i>
Rhizomelic chondrodysplasia punctata	<i>PEX7</i>

<sup>a</sup>Other forms of epidermolysis bullosa are due to Lamin gene mutations (laminopathies).

<sup>b</sup>Diseases under the section of mitochondriopathies relate to both unfolded/misfolded proteins (proteinopathies) and indirect, cumulative effects of increased oxidative stress, but not to direct effects on mitochondria.

ALS, Amyotrophic lateral sclerosis.

histones, directly regulating transcription activity or repression, depending on each specific combination. For transcription to take place, chromatin harboring DNA elements such as gene promoters and enhancers needs to be less condensed, thus accessible for transcription factor (TF) binding. The promoter is a limited segment of the 5' region of the gene displaying binding sites for general TFs. Promoters of human protein-coding genes are recognized by TF functioning with RNA polymerase II (POLII). Therefore they have the denomination of RNA POLII transcription factors (TFII), known by the acronyms TFIIA to TFIH. Historically, promoters for RNA POLII have originally been identified as the TATA box, a short motif with conserved base core thymidine–adenine–thymidine–adenine recognized by TATA-binding protein (TBP), one TFIID subunit. It is now acknowledged that nearly 50% of human protein-coding genes have one TATA motif in their promoters. To compensate for the lack of TATA representation in other promoters, additional motifs have evolved as binding sites for TBP or other TFII subunits, as part of TFIID or TFIIB. Therefore the core promoter of protein-coding genes has now been recognized as a stretch of nearly 100 base pairs (bps) comprehending the transcription start site (TSS, enumerated as +1 for the first transcribed nucleotide), up- and downstream sequences, displaying a variable combination of two or more motifs for TFII binding. In fact, while most promoter motifs locate upstream the TSS, others such as the initiator and downstream elements are positioned respectively on the TSS and downstream to it. Thus the core promoter in general encompasses bases –45 to +30, enumerated in regard to the TSS (Fig. 2.2). Transcription initiation involves interaction of TFIID with the core promoter. TFIIB and other TFIs sequentially establish protein–protein contacts among them, assembling a multiprotein complex on the core promoter, which allows for DNA local denaturation and RNA POLII association immediately downstream. TFIH phosphorylates POLII, triggering transcription initiation and recruiting RNA processing factors, such as capping, splicing, and polyadenylation factors, to POLII C-terminal domain. Therefore every cell has constitutive expression of TFIs as they are fundamental proteins for positioning RNA POLII to initiate transcription of every protein-coding gene.

Transcription activation by the core promoter is responsible for basal levels of pre-mRNA production, known as constitutive transcription. In most cases, other DNA elements are bound by tissue-specific TFs (sTFs), exponentially enhancing transcription levels. These DNA elements may be positioned far away from the promoter, even 30 kb upstream or downstream of the core promoter (Fig. 2.2). They are activated if bound by a protein factor with a tissue-specific expression pattern. Therefore activation will only take place if that *trans*-acting protein factor is expressed in the respective cell type. Likewise, DNA silencers with varying distribution in respect to the promoter may bind protein factors with negative effect on transcription, known as transcription repressors. As enhancers and silencers may have variable and distant localizations relative to the promoter position, a large complex should be assembled to communicate TFII- and POLII-associated promoter to a distant enhancer/silencer bound to activation/repression factor. Mediator is a large multimeric complex composed of MED proteins that assemble as four smaller complexes or modules. Mediator head module associates with TFII–POLII complex, whereas the tail module interacts with distant tissue sTF bound to enhancer or silencer. The middle module bridges the other modules, looping the chromatin between both DNA elements, enhancer/silencer and promoter (Fig. 2.2). As the tail module needs to interact with different TFs depending on cell type, its composition may vary. The fourth is the mediator kinase module that interacts with other modules when POLII is unavailable thus maintaining the mediator complex in a closed inactivated conformation. Mutations in genes coding for MED units have been reported causing genetic diseases (see Genetic case 3B).



*Genetic case 3B:* A 19-year-old man presented with developmental delay and learning disabilities since childhood. Family history was unremarkable, and no parental consanguinity was reported. On physical examination the patient presented with hypotonia, high stature, prominent forehead, long narrow face, low-set ears, small mandible, and long arms and fingers. Brain and spine MRI, ophthalmologic and cardiologic exams were normal. Karyotype was 46, XY, negative for fragile X genetic testing. Exome sequencing disclosed a mutation in the X-linked gene *MED12*, consistent with Opitz–Kaveggia/Lujan–Fryns syndromes.

### 2.3.2 Cajal bodies, speckles and pre-mRNA processing

In the nucleoplasm, there are subnuclear structures locally concentrating ribonucleoprotein factors involved in maintenance of telomeres, gene transcription, pre-mRNA and rRNA precursor processing, and other aspects of ribosome biogenesis. All structures are unbound of membrane and in general organized around small or large noncoding RNAs. There are two mostly prominent nuclear bodies; the nucleolus (see later) and Cajal bodies. Cajal bodies are storage and distribution centers of ribonucleoprotein complexes to the nucleolus and speckles. Their number in a cell nucleus is limited and variable. These nuclear foci contain a vast variety of proteins associated with small nuclear RNA (snRNA) and small nucleolar RNAs (snoRNA), essential for pre-mRNA splicing and ribosome biogenesis, respectively.

As an overview, snRNAs are transcribed by RNA POLII, processed, and exported from the nucleus. In the cytoplasm, different proteins, including the survival of motor neuron 1 (SMN), associate with snRNA assembling small nuclear ribonucleoprotein particles (snRNP), which are modified and imported to the nucleus, reaching Cajal bodies. SnRNPs are further matured in Cajal bodies and then move to nuclear speckles. This flow of newly synthesized snRNPs is required for Cajal body integrity. Similar situation occurs for the maturation of snoRNPs and distribution of these nucleolar complexes. An example of dysfunction in this pathway is the autosomal recessive spinal muscular atrophy, which is due to mutation in the *SMN1* gene (see Genetic case 3C).

*Genetic case 3C:* A 10-year-old boy with spinal muscular atrophy confirmed by the finding of a biallelic mutation in the *SMN1* gene has a pediatrics consult for annual growth and development evaluation. Neurological features of the disease have been bilateral proximal muscle weakness and atrophy. Parents show concerns about snoring, frequent abdominal breathing pattern, increasing fatigue, and feeding difficulties.

Numerous nuclear bodies reside in the interchromatinic space. Among them lie speckles or interchromatinic granule clusters, where splicing factors finish their maturation and assembly, remaining readily available for perichromatin fibrils, when transcription is initiated (Fig. 2.2). In perichromatin fibrils, actively transcribing genes are associated with nascent pre-mRNAs and RNA POLII C-terminus domain harboring pre-mRNA processing factors in physical association. This interaction is only possible after transcription has been initiated, as that RNA POLII domain becomes phosphorylated in specific serines from heptad repeats. Processing factors are thus recruited from speckles to RNA POLII C-terminus domain upon transcription start. This close association between processing effectors and nascent pre-mRNA elicits cotranscriptional pre-mRNA processing, which comprehends 5'-capping, pre-mRNA splicing, and 3'-polyadenylation.

As soon as the 5'-end of pre-mRNA is synthesized by RNA POLII, capping effectors are activated. 5'-capping consists of addition of a 7-methyl-guanosine to the 5'-end ribonucleotide by atypical 5'-5' triphosphate bond. Specific nuclear proteins are then associated with the 5'-cap. Pre-mRNA splicing occurs as long as a series of exon–intron–exon appears in the nascent RNA molecule. Apart from the canonical intron motifs that signal to the spliceosome (Fig. 2.2), specific proteins, such as serine- and arginine-rich (SR family) proteins or heterogeneous nuclear proteins (hnRNP family), tend to associate with nascent RNA exons and introns, respectively. Coating the RNA by exon- or intron-specific proteins strengthens exon and intron definitions on the pre-mRNA for recognition by the nuclear splicing machine, the spliceosome. The spliceosome is composed of five snRNPs, named U1, U2, U4, U5, and U6, each containing one molecule of the respective snRNA and a large variety of proteins. SnRNP U1 and U2 bind to intron 5'-end and the branch point-containing segment, respectively, by specific RNA–RNA and RNA–protein interactions. A rapid rearrangement takes place permitting the association of snRNPs U4–U5–U6. U1 and U4 release allows for U6 and U2 snRNA base-pairing, bringing intron ends close to each other. It also permits the assembly of the spliceosome catalytic site, composed of U5 and U6 snRNPs.

Splicing itself consists of two sequential transesterification reactions. The first reaction frees the 5' exon from downstream intron and consists of a nucleophilic attack of the adenine from intron branching point by intron 5'-end nucleotide (a guanine), making an unusual covalent bond between intron 5' end and adenine 2' carbon hydroxyl. The second transesterification reaction is the phosphodiester bond between the two exons. Importantly, splicing catalysis is in part performed by snRNA ribozyme activity. After the second transesterification reaction, a protein complex is deposited at each exon–exon junction, the exon junction complex (EJC), which is important for mRNA quality control (see later). Constitutive splicing consecutively maintains all exons in the mature mRNA. Alternatively, according to exon and intron lengths, binding strength and coverage by SR and hnRNP proteins, the spliceosome can alter the mRNA output in the following ways: exon skipping, choice of mutually exclusive exons, usage of alternative 5' or 3' intron donor or acceptor sites, or intron retention (Box 2.1 and Fig. 2.2).

Finally, as RNA POLII transcription proceeds beyond the translation termination codon, the 3'-untranslated region of pre-mRNA is productively synthesized until a polyadenylation motif is transcribed. Its recognition by polyadenylation protein effectors leads to pre-mRNA cleavage nearly 30 nucleotides downstream the motif, decreasing processive transcription by RNA POLII and possibly reducing its affinity to the DNA, an accepted model of transcription termination in eukaryotes. Upon cleavage, poly-A polymerase synthesizes the poly-A tail, which typically contains an average of 200 nucleotides.

### 2.3.3 Nucleolus

Ribosome is a huge ribonucleoprotein particle composed of small and large subunits, with Svedberg (S) coefficients respectively of 40S and 60S in eukaryotes. Ribosomal RNAs (18S in the small subunit and 23S, 5.8S, and 5S in the large subunit) fold into self-complementary helices and associate with 33 or 49 proteins to assemble each subunit, respectively. There are five loci for rRNA genes in the human genome, known as nucleolar organizing regions. In telophase, daughter cell promoters for rRNA genes initiate transcription by RNA polymerase I. Thus by M/G1 phase transition, 10 loci are synthesizing rRNA at high rates. Ribosomal RNAs are posttranscriptionally modified by the aid of snoRNA. Ribosomal proteins associate with rRNA, and assemble a large 90S ribosome precursor. Its nuclear processing leads to the formation of both ribosome subunits 40S and 60S, which are exported to the cytoplasm. The *TCOF1* gene encodes treacle, a protein involved in different aspects of ribosome biogenesis in neural crest cells. Mutations in *TCOF1* cause Treacher Collins syndrome, an example of ribosomopathy (see Genetic case 3D, Table 2.1).

---

*Genetic case 3D:* A female toddler has been assisted by the oral and maxillofacial surgeon team, as she was prenatally diagnosed with Treacher Collins syndrome, a craniofacial dysostosis disorder. Besides congenital skeleton abnormalities, prenatal diagnosis was confirmed by the detection of a deletion in the *TCOF1* gene. Along the first year of her life, she had been frequently at hospital for airway management. At age of 16 months, she has been in postoperative period of cleft palate repair and followed up by otologists as she has conductive hearing loss. She has been referred to speech therapy and will be monitored along her growth as bone graft surgeries for mandible and zygomatic bone correction are predicted for ages 4–5.

---

### 2.3.4 Nuclear envelope and mRNA quality control

The nuclear envelope consists of two membranes perforated by a few thousand pores. The small GTPase ran (Table 2.2) coordinates a two-way flow of molecules through the pore, allowing passage of cargoes by association with carrier molecules. This interaction takes place due to the existence of motifs in cargo proteins for nuclear entry or exit, respectively, nuclear localization and export signals.

Mature RNA messages are transported through the envelope pore from the nucleus to the cytoplasm. In the adjacency of the nuclear pore, part of nuclear proteins bound to the mRNA is switched for their cytoplasmic paralogs or isoforms. EJC proteins, however, remain bound to mRNA until a pioneer ribosome translation round shifts them. In that single ribosome round on the mRNA, all EJC should be removed unless there is a premature translation stop codon. Messages with premature stop codon may arise due to mutations, alternative splicing or RNA POLII errors. A cell surveillance mechanism has evolved where the permanence of EJC on mRNA after the pioneer translation round signals to mRNA decay/degradation systems. Consequently, this process (nonsense-mediated



**TABLE 2.2** Characteristic examples of GTPase protein family members and respective regulated biological processes.

GTPase gene acronym <sup>a</sup>	Regulated biological process
<b>I. Guanine nucleotide-binding (G) protein subunit alpha family (from trimeric GTPases)</b>	
<i>GNAS</i> (from G protein alpha group S—stimulatory)	Signal transduction by cAMP
<i>GNAI1</i> (from G protein alpha group I—inhibitory)	Signal transduction by cAMP
<i>GNAQ</i> (from G protein alpha group Q)	Signal transduction by phospholipase C
<i>GNA12</i> (from G protein alpha group 12/13)	Signal transduction by phospholipases C/D
<b>II. GTP translation factor family</b>	
<i>EIF2S3</i> (Eukaryotic translation initiation factor 2 gamma)	Translation initiation
<i>EEF1A</i> (Eukaryotic translation elongation factor 1 alpha)	Translation elongation
<i>EEF2</i> (EF2— Eukaryotic Translation Elongation Factor 2)	Translation elongation
<i>ETF1</i> (Eukaryotic Translation Termination Factor 1)	Translation termination
<b>III. Dynamin GTPase superfamily</b>	
<i>DNM1</i> (Dynamin-1, classical dynamin)	Vesicular trafficking
<i>DNM1L</i> (Dynamin-like protein)	Mitochondria and peroxisome fission
<i>OPA1</i> (Mitochondrial dynamin-like protein, OPA proteins)	Mitochondria fusion and apoptosis regulation
<i>MX1</i> (MX dynamin-like protein, MX proteins)	Interferon-induced cellular antiviral response
<i>MFN1</i> (Mitofusins)	Mitochondrial fusion
<i>ATL1</i> (Atlastins)	Endoplasmic reticulum membrane fusion
<b>IV. Ras monomeric small GTPase superfamily<sup>b</sup></b>	
Ras family	
<i>HRAS</i>	Cell proliferation
<i>KRAS</i>	Cell proliferation
<i>NRAS</i>	Cell proliferation
<i>RAP1A</i>	Counteracts Ras mitogenic function
<i>RHEB</i>	Cell growth
Rho family	
<i>CDC42</i>	Actin remodeling (filopodium formation)
<i>RAC1</i>	Actin remodeling (ruffle formation)
<i>RHOA</i>	Actin remodeling (stress fiber formation)
Ran family	
<i>RAN</i>	Nucleo-cytoplasmic shuttling
Rab family	
<i>RAB1A</i>	Vesicle transport from ER to Golgi apparatus
<i>RAB3A</i>	Exocytosis
ARF (ADP ribosylation factor) family	
<i>ARF1</i>	ER and Golgi apparatus vesicle budding
<i>ARL8</i>	Lysosomal biogenesis and motility
<i>ARL13A</i>	Cilium formation and maintenance
Rab-like Family	
<i>RABL2A</i> and <i>IFT22</i>	Intraflagellar molecule transport

(Continued)

TABLE 2.2 (Continued)

GTPase gene acronym <sup>a</sup>	Regulated biological process
Mitochondrial Rho proteins (MIRO) atypical GTPases	
<i>RHOT1</i> (Miro-1 protein) and <i>RHOT2</i> (Miro-2 protein)	Mitochondrial trafficking

<sup>a</sup>If the protein name is different from the gene name, it is indicated in parenthesis, as well as its subgroup classification, when necessary.  
<sup>b</sup>There are more than 150 human genes encoding members of the Ras small GTPase superfamily available at HUGO (Human Gene Nomenclature Committee) <https://www.genenames.org/cgi-bin/genefamilies/set/358> (accessed in December 2017).  
ER, Endoplasmic reticulum; GTP, guanosine 5'-triphosphate.

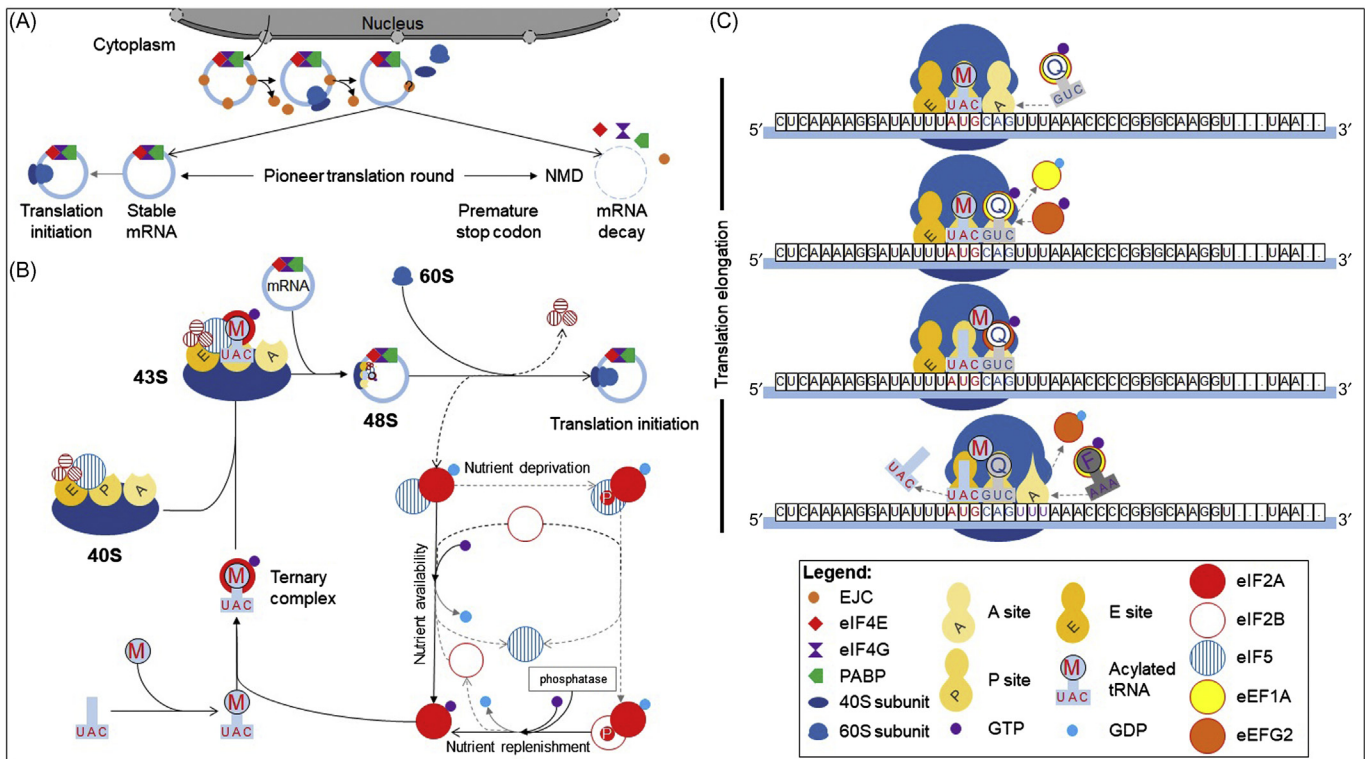
mRNA decay) may destroy many mRNAs that would be translated into truncated proteins, thus reducing the stress of unfolded proteins in the endoplasmic reticulum (ER) and cytoplasm (Fig. 2.3).  
The cell employs three major ways to destroy mRNA. All of them ultimately render the mRNA vulnerable to either 5'-to-3' or 3'-to-5' exonucleases. Decapping, the loss of mRNA 5' modified ribonucleotide, and poly-A tail deadenylation are most commonly employed in mammalian cells. Decapping activates 5'-to-3' exonucleases to successively cleave phosphodiester bonds from the 5' end. Along the mRNA life, it progressively loses adenines from its long poly-A tail. When the 3'-tail length shortens to a low threshold near 30 bases, 3'-to-5' exonucleases rapidly digest the molecule. Alternatively, in specific cases, the third manner to initiate RNA cleavage is by endonucleases, enzymes that cleave phosphodiester bond between ribonucleotides internally located in the RNA strand. Consequently, the 5'-end product will be substrate for 3'-to-5' exonucleases, whereas the 3'-end by-product will undergo 5'-to-3' exonucleolytic cleavages.

## 2.4 Protein synthesis

When mRNA reaches the cytoplasm, its 5'-cap becomes associated with eukaryotic translation initiation factor 4E (EIF4E), and the poly-A tail interacts with units of cytoplasmic polyadenylate-binding protein (PABP). A circularized configuration is adopted as eukaryotic translation initiation factor 4G (EIF4G) bridges mRNA 5'- and 3'-ends by binding to both proteins EIF4E and PABP (Fig. 2.3). The circular mRNA configuration is a strong requisite for translation initiation. Intervention on this protranslational configuration may take place if intervening small RNAs such as microRNA or additional proteins as translational repression factors bind the mRNA. In these cases, mRNA is not translated but may remain stable in the cytoplasm in a dormant state. These interventions—RNA interference or protein-mediated translation repression—are effective ways to downregulate protein synthesis (Box 2.1). Alternatively, similar strategies can trigger cytoplasm mRNA degradation by the mechanisms previously presented.  
Protein synthesis is unlikely to take place if amino acids are not available. Therefore nutrient withdrawal is a strong, general signal to overall shutting down protein synthesis. Starvation has two main mechanistic effects on protein synthesis. The first one is on mRNA configuration by competition for EIF4E binding. Along the phylogenetic scale, DNA encoding the EIF4E-binding motif, as in the *EIF4G1* gene, has arisen in other protein-coding genes. One of these genes encodes eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1) that, similarly to EIF4G, is able to interact with EIF4E. Amino acid availability favors EIF4G binding to EIF4E, as it triggers 4E-BP1 phosphorylation, which inversely relates to its ability to bind EIF4E. Upstream signaling to 4E-BP1 depends on mTOR complex 1, a conserved, regulatory kinase, upregulated when the ratio between AMP and ATP is low. Conversely, high AMP/ATP ratio indicates nutrient starvation thus leads to mTOR inactivation and increases binding of dephosphorylated 4E-BP1 to EIF4E, decreasing the probability of mRNA translation initiation. Upon cell replenishing with nutrients, AMP/ATP ratio reduces, there is consequent activation of mTOR, which phosphorylates 4E-BP1, lowering its affinity to EIF4E. The latter then preferentially binds to EIF4G resuming the mRNA protranslational configuration.

### 2.4.1 Ribosomes and mRNA translation

Apart from mRNA, transfer RNAs (tRNA) and ribosome are functional components essential for protein synthesis. tRNA are short RNA molecules folded into hairpins. The middle hairpin contains the anticodon in its loop, whereas in the 3' end of the tRNA molecule a specific amino acid may be covalently bound.



**FIGURE 2.3** Major highlights on mRNA translation initiation regulation by energy and amino acid availability, and on translation elongation. (A) Upon nuclear export of mRNA, a pioneer translation round takes place in the nuclear pore boundaries, removing all EJC. NMD may be triggered and mRNA thus degraded if an EJC is downstream to a premature stop codon. Otherwise, mRNA associated with EIF4E, EIF4G, and PABP, free of EJC, can be stably distributed in the cytoplasm and have translation initiated. (B) Formation of the 43S complex under favorable nutrition conditions leads to translation initiation. For that, tRNA needs to be acylated, in particular tRNA for the start codon should carry methionine (M). Then, the ternary complex of Met-acylated tRNA and GTP-bound EIF2A is able to interact with 40S ribosome subunit associated with EIF5 and other EIF (small, red striped circles), generating the protranslational 43S complex. During translation initiation, EIF5 functions as EIF2-GTP GAP. EIFs are released for translation initiation and ribosome assembly. It is possible that upon nutrient deprivation EIF5 functions as EIF2-GDP GDI indirectly rendering EIF2A phosphorylation, then being replaced by EIF2B. Nutrient replenishment elicits EIF2A dephosphorylation and allows EIF2B to exchange GDP by GTP. (C) Translation elongation initiates with Met-tRNA<sup>Met</sup> on the P site. Acylated tRNA (in this case carrying glutamine, Q) is positioned on the A site by aid of the GTPase eEF1A. After ribosome catalysis of peptide bond formation, eEF2G shifts the ribosome allowing novel acylated tRNA to be positioned in A site with eEF1A. In both cases, there should be GTP hydrolysis. Deacylated tRNA on E site may be released to the cytosol. EIF4E, Eukaryotic translation initiation factor 4E; EIF4G, eukaryotic translation initiation factor 4G; EJC, exon junction complexes; GDI, GDP dissociation inhibitor; GTP, guanosine 5'-triphosphate; NMD, nonsense-mediated mRNA decay; PABP, polyadenylate-binding protein; tRNA, transfer RNAs.

Aminoacyl-tRNA synthetases are amino acid-specific enzymes that catalyze the attachment (acylation) of amino acid to tRNA 3' end. After a tRNA has provided its amino acid for the synthesis of a protein, it leaves the ribosome deacylated. If the cell environment is energetically favorable, tRNA acylation takes place in the cytosol by specific aminoacyl-tRNA-synthetases. However, if the cell is deprived of nutrients, the probability of tRNA acylation is considerably low. Therefore the second mechanistic effect of nutrient starvation on protein synthesis relates to tRNA acylation regulation by amino acid availability, more specifically the acylation of tRNA<sup>Met</sup>, which recognizes the translation initiation codon (AUG). If plenty of amino acids is available, methionine is attached to tRNA<sup>Met</sup> (Met-tRNA<sup>Met</sup>), and this complex is likely to interact with GTP-bound eukaryotic translation initiation factor 2A (EIF2A), constituting the ternary complex (Fig. 2.3; EIF2-GTP-Met-tRNA<sup>Met</sup>).

Two ribosome subunits, 40S and 60S, compose the 80S ribosome only during translation. If unbound to mRNA, they occur separately in the cytosol. The formation of the ternary complex is conditional for the delivery of acylated tRNA (Met-tRNA<sup>Met</sup>) to the ribosome 40S subunit. When this subunit combines with the ternary complex and other specific eukaryotic translation initiation factors, the larger assembled complex becomes the 43S translation preinitiation complex. The configuration of ribosome small subunit as 43S is thus a strong requisite for translation initiation (Fig. 2.3). Together, amino acid availability and cytosol AMP/ATP ratio control GTP loading to EIF2A through another EIF2 subunit that functions as GTP exchange factor (GEF). In addition,

persisting starvation leads to EIF2A phosphorylation and binding to GDP dissociation inhibitor maintaining translation initiation repression (see [Section 2.8](#) and [Fig. 2.6](#)). Consequently, if the cell is starved, EIF2A is not bound to GTP, and the ternary and 43S complexes are not formed. Therefore cell nutrition regulates the 43S complex assembly by interfering in tRNA<sup>Met</sup> acylation and the association between GTP and EIF2A ([Fig. 2.3](#)).

Protein synthesis can be divided into three phases: initiation, elongation, and termination. As presented, the formation of mRNA protranslational configuration and the 43S ribosome translation preinitiation complex are two main requisites for translation initiation. The 43S preinitiation complex associates with EIF4E-bound mRNA 5'-cap, and scans the 5'-untranslated region for the translation initiation code. When Met-tRNA<sup>Met</sup> anticodon base pairs with the initiator code AUG, EIF2A hydrolyzes its bound GTP. Then translation initiator factors, including EIF2A, dissociate from the ribosome small subunit, allowing it to bind to the large subunit, assembling the ribosome. When translation initiation finishes, Met-tRNA<sup>Met</sup> is positioned at ribosome P site ([Fig. 2.3](#)).

Translation elongation phase is the protein synthesis itself. For each protein bond the following sequence of events must occur: (1) binding of an amino acid-tRNA to the ribosome A site; (2) proofreading of base pairing between mRNA codon and tRNA anticodon by the ribosome decoding center; (3) peptide bond formation; (4) ribosome translocation, shifting the mRNA by one codon, moving the nascent peptide to the P site, thus freeing the A site; and (5) release of deacylated tRNA from the ribosome E site to the cytosol. At this point, in the ribosome P site, the last tRNA contains the nascent peptide attached to its 3' end (referred to as peptidyl-tRNA). The sequence of events will repeat for every coming amino acid-tRNA. In this sequence, two aspects should be highlighted. The first one is the need of two distinct GTPases for translation elongation; GTPase eEF1A charges amino acid-tRNA to the A site, and GTPase eEF2, which consecutively occupies the A site, transfers the peptidyl-tRNA to the P site, and slides mRNA by three bases exposing the following code in the A site. Both functions are mediated by GTP hydrolysis ([Table 2.2](#) and [Fig. 2.3](#)). The second remark is that rRNA acting as ribozyme is responsible for peptide bond formation catalysis in the peptidyl transferase center of the ribosome large subunit.

Among 64 three-base codon possibilities, three do not encode amino acids but correspond to translation termination (stop) codons (UAA, UAG, or UGA). No tRNA recognizes a stop codon. Instead, when a termination codon moves into ribosome A site, a eukaryotic release factor (eukaryotic translation termination factor, [Table 2.2](#)) recognizes it and, due to its GTPase activity, induces the ribosome peptidyl transferase center to hydrolyze the peptidyl-tRNA, releasing the peptide. A conformational change allows water entry between 40S and 60S ribosome subunits, dissociating and releasing them in the cytosol.

## 2.5 Vesicular trafficking: the secretory and endocytic pathways

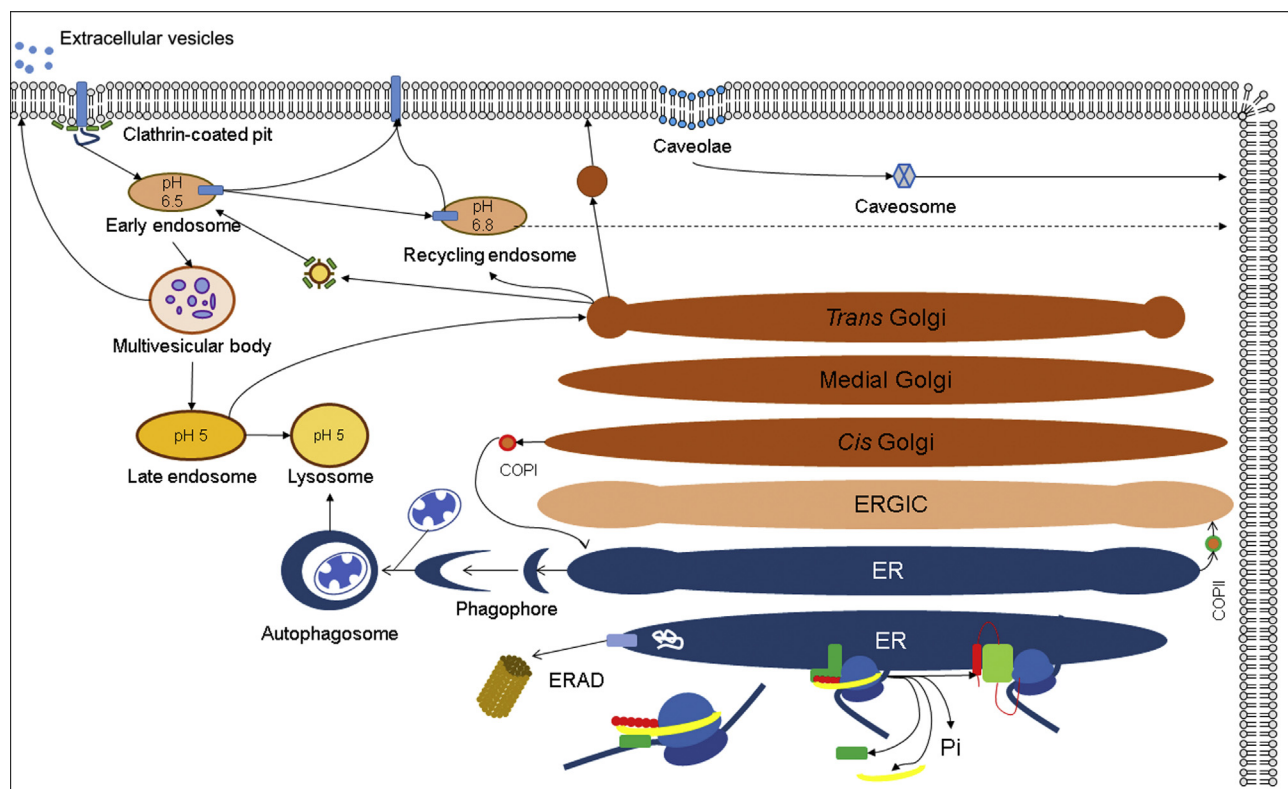
Transport of molecules or macromolecular complexes between organelles involves vesicle budding from one membrane-bound organelle and fusion with the membrane from the destination organelle, a process called vesicular trafficking. There are two major pathways in this process, the secretory and the endocytic pathways. The secretory pathway coordinates organelle biosynthesis and cell secretion and targets proteins to their home organelle. Although it works in predominantly vectorial direction, from the ER to the plasma membrane or lysosomes, retrograde transport is necessary for the retrieval of organelle resident proteins. Most membrane lipids are synthesized in the smooth ER membrane (not presented in this chapter). The rough ER (RER) synthesizes membrane proteins and soluble proteins destined to the secretory pathway, the endosome, and lysosomes.

Nearly one-third of proteins encoded by the human genome will be targeted to the endomembrane system thus directed to the RER. During mRNA translation by cytosolic ribosome, the nascent peptide may display at its N-terminus a series of 15–35 amino acids with a hydrophobic core, termed signal sequence or signal peptide. Signal sequences may be found in integral membrane proteins of the plasma membrane; membrane from organelles of the secretory pathway; or in soluble proteins residing in the lumen of those organelles. Once the protein signal sequence is synthesized by cytosolic ribosomes, it is recognized by a ribonucleoprotein particle (signal recognition particle, SRP) pausing translation; then the complex is forwarded to the ER, where protein synthesis is resumed. If the signal sequence is absent, hydrophobic domains in any other location in the protein sequence may signal to and interact with the SRP, sorting it to the ER.

SRP ribonucleoprotein complex is composed of a 300-nucleotide RNA (7SL RNA) on which six proteins are associated. It is a large elongated complex that wraps the large ribosomal subunit. One end binds to the nascent protein signal sequence, and the other end binds the elongation factor binding site (A) in the ribosome, halting protein synthesis as soon as the signal sequence emerges from the ribosome. There are two major protein



complexes in the RER membrane important for cotranslational import of transmembrane proteins. The SRP receptor is a heterodimer having one integral membrane protein and a soluble unit that directly interacts with SRP-bound ribosome. Receptor binding attaches SRP-ribosome to the RER by SRP receptor heterodimerization, which hydrolyzes receptor- and SRP-bound GTPs causing their dissociation, and transferring the ribosome to the second protein complex, a multimeric protein-conducting channel also known as translocon. The translocon provides a high-affinity binding site for the complex of ribosome, the translationally arrested mRNA, and the nascent peptide chain harboring the signal sequence. This interaction leads the ribosome to resume protein synthesis, as the A site becomes available. Ribosome–translocon association provides a path for nascent protein transfer from the peptidyl transferase center to the RER lumen. It is accepted that as the peptide N-terminus is cytosolic, the transmembrane domain (signal sequence) spans the ER membrane, and the following segment loops in the RER lumen in continuity with the translocon, which is an aqueous channel (Fig. 2.4). Translation elongation pushes the nascent peptide through the channel using the energy from protein synthesis. In multipass transmembrane proteins, the second transmembrane domain should be laterally transferred from the translocon to the membrane; and translation continues in the cytosol. If there is a third transmembrane domain in the protein, the process of cotranslation import to the RER starts over. The signal peptide may be retained in the mature protein, cleaved in the RER releasing the protein to its lumen, or cleaved and retained in the RER membrane by a second transmembrane domain.



**FIGURE 2.4** The secretory pathway and endocytic pathways. On the bottom of the figure, the synthesis of a protein carrying a sequence signal (red) is highlighted by the association of the ribosome to the ER. This takes place due to the interaction of SRP (yellow) which consecutively interacts with soluble unit of SRP receptor (green) directing the complex to the ER membrane dimerizing the receptor. Hydrolyses of GTP molecules associated with SRP and SR shifts the transmembrane domain to the translocon (light green rectangle) and resumes protein synthesis. From the ER, cytosol translocation of unfolded proteins forwards them to the proteasome; membrane fragmentation leads to formation of the phagophore maturing into autophagosome (indicated by mitophagy). COPII- and COPI-coated vesicles indicate respectively forward and retrograde ER–Golgi transport. Soluble proteins in association with mannose-6-phosphate receptor leave *trans* Golgi in clathrin-coated vesicles and pass through early and late endosomes before arriving in lysosomes. Endocytosis started in clathrin-coated pits directs cargo to early endosome, which can direct back to the plasma membrane, and to it through recycling endosomes or forward cargo to late endosome passing by multivesicular bodies. Along the pathway of clathrin-mediated endocytosis, successively lower pH is important for cargo release and degradation. Slight increase in pH is relevant for cargo recycling. Multivesicular bodies are also responsible for exocytosis of exosome. Autophagosome and late endosome will be targeted to the lysosome for content hydrolysis. Caveolae-initiated endocytosis is important in transcellular transport. *COP*, Coat protein complex; *ER*, endoplasmic reticulum; *SRP*, signal recognition particle.

In the ER lumen, different proteins may associate with and modify the nascent polypeptide. Chaperones promote folding of polypeptide chains into three-dimensional conformations and assembly of multisubunit proteins. In addition, protein disulfide isomerase catalyzes disulfide bond formation assisting protein folding. Covalent modifications of nascent polypeptides (posttranslational modifications) occurring in the ER lumen include glycosylation [e.g., addition to asparagine (N-linked glycosylation) and partially trimming of carbohydrate branched chains], and protein cleavage and transfer to membrane-anchored glycolipids (GPI tails).

Different neurodegenerative diseases, although clinically and etiologically distinct, may share a common pathological trait—the abnormal aggregation of misfolded proteins leading to progressive ER stress and thus been classified as proteinopathies (Table 2.1). ER stress triggers a coordinated biochemical response known as the unfolded protein response (UPR), preventing the export of unfolded, dysfunctional proteins into the secretory pathway. UPR is activated by unfolded proteins in the ER lumen, recognized by specific ER transmembrane proteins functioning as unfolding sensors. If activated, specific membrane sensors elicit cytoplasmic and nuclear signaling responses ultimately leading to one of three major UPR outcomes: (1) translocation of unfolded proteins from the ER lumen to the cytosol for degradation (ER-associated degradation, Fig. 2.4); (2) overall arrest of protein synthesis initiation through the inhibition of EIF2A by phosphorylation (Section 2.4); or (3) tentative repair of defective proteins by the mobilization of ER-resident protein chaperones.

Properly folded proteins destined to move ahead in the secretory pathway are transported to the Golgi apparatus. The recognition of short specific amino acid signatures within the protein (motif) is necessary for its transport to the ER–Golgi intermediate compartment and then to Golgi employing the vesicular pathway. Proteins displaying the ER-retention motif are retrieved from Golgi to the ER by retrograde transport. There are four major protein classes necessary for vesicular trafficking: small GTPases (Table 2.2, Section 2.8), membrane coating molecules, tethering factors, and fusion proteins. Vesicle budding is activated by small GTPases. Bound to GTP, small GTPases as ARF1 or rab are tightly associated with the cytosolic face of the membrane and recruit effector proteins, comprehending soluble coat protein complexes I (COPI) or II (COPII), two distinct multimeric complexes. COPII-coated vesicles bud from the ER in anterograde transport, whereas COPI tags vesicles budding from Golgi membrane aiming to retrograde transport to the ER (Fig. 2.4). The importance of retrograde transport of ER-resident proteins is illustrated by heterozygous loss-of-function mutations in the *ARCN1* gene, which encodes one COPI protomer, leading to growth and developmental delay syndrome (see Genetic case 5A), and in vitro ER stress. Membrane coating by coatomer complexes I or II takes place sequentially, and the dynamics of the process depends on the activity status of the respective small GTPase. Consequently, coat assembly and disassembly are regulated processes. Once the vesicle buds and pinches off the donor membrane, coatomer disassembly is favorable. Uncoated vesicle interaction with acceptor membrane depends on rab small GTPases, as well as tethering and fusion proteins. Tethering factors are transmembrane proteins originally incorporated in the vesicle from the donor membrane. They extend a long, rod domain in the cytosol, which coils with the tethering factor from the acceptor membrane, establishing the first contact between the two membranes. Membrane fusion takes place when SNARE (soluble attachment receptor) proteins from donor (e.g., vesicle, thus named v-SNARE) and acceptor membranes (as targeted destination, hence t-SNARE) interact, forming extensive coiled coils that progressively bring vesicle and acceptor membranes closer, leading to hydrolysis of rab-bound GTP.

---

*Genetic case 5A:* A 9-year-old girl was referred to the genetics clinics due to developmental delay. The acquisition of language skills was slower than that of her peers, and she was as yet unable to read. Physical examination revealed microcephaly, micrognathia, short stature, and rhizomelic shortening. Exome sequencing revealed heterozygous mutation in the *ARCN1* gene. Literature review disclosed four other cases with similar genetic and clinical features.

---

The Golgi apparatus (Golgi complex) is a stack of flattened cisternae and associated vesicles. Due to differential molecular composition and specializations, Golgi cisternae are named *cis* Golgi if proximal to the ER, *trans* Golgi when closer to the plasma membrane, and the medial compartment lies between them. Golgi is involved in further glycoprotein modifications, major steps in proteoglycan synthesis for constitution of the extracellular matrix, and sphingolipid biosynthesis. In the Golgi complex, additional protein modifications take place, as pro-hormones may undergo cleavage, and proteins are loaded with metal or lipids, respectively as metal transport proteins and enzyme cofactors, or lipoproteins. As example, the Golgi-resident ATPase ATP7B is a copper-transporting ATPase fundamental for copper homeostasis. In hepatocyte *trans* Golgi, ATP7B loads copper onto specific proteins (e.g., ceruloplasmin) and forwards them to the plasma membrane basolateral domain for

secretion. Excessive copper directs ATP7B from TGN to hepatocyte apical domain where it mediates secretion of free copper to the bile. Homozygous mutations in the *ATP7B* gene cause Wilson disease, a genetic disorder of copper metabolism (see Genetic case 5B).

---

*Genetic case 5B:* A young adult male was admitted with history of asymmetrical tremor and changes in behavior. Evaluation showed mild hepatomegaly and Kayser–Fleischer rings upon slit-lamp examination. Serum ceruloplasmin was reduced, 24-hour urine copper considerably raised, and liver biopsy revealed a 20-fold increase in copper deposition. Wilson disease is clinically heterogeneous, ranging in the young adult from asymptomatic to chronic liver disease, with or without cirrhosis, or fulminant hepatic failure, and varying neurologic and psychiatric manifestations.

---

The Golgi apparatus is also the major protein-sorting hub in the secretory pathway. Cargo sorted from the *trans* Golgi can be distributed to the plasma membrane and cell exterior as well as to different intracellular compartments, such as the endosomal/lysosomal system, specialized secretory organelles or granules. Golgi constitutive sorting pathway is composed of tubular transport carriers that bud from the *trans* Golgi to the cell surface, apparently in uncoated manner. Transmembrane proteins sorted to the plasma membrane are surrounded by bilayer domains enriched in sphingolipids and cholesterol. Tubule extension from the *trans* Golgi membrane is activated by phosphatidylinositol kinases, whereas tubule severing with vesicle formation is mediated by dynamin (GTPase family, Table 2.1), characterizing exocytosis as active transport. In nonconstitutive exocytosis, calcium is an important signaling molecule, such as in exocytosis of neurotransmitters from presynaptic vesicles. These vesicles contain proteins playing roles in regulation of exocytosis triggered by calcium. For instance, otoferlin is essential for auditory hair cell synchronous exocytosis as demonstrated by mutations that in homozygosis cause sensorineural hearing loss (see Genetic case 5C).

---

*Genetic case 5C:* An 18-month-old male toddler has profound, bilateral hearing impairment due to homozygous mutations in the *OTOF* gene, which codes for otoferlin. Parents take the child for a consult in the otology clinics, seeking information on cochlear implant.

---

Mechanisms to direct hydrolytic enzymes toward lysosomes are apart from the constitutive secretory pathway to the plasma membrane. Lysosomal hydrolytic enzymes are glycoproteins that in Golgi complex acquire a phosphate on carbon 6 of a mannose. Mannose-6-phosphate binds mannose-6-phosphate receptors in the *trans* Golgi lumen, a targeting signal to the endosome/lysosomal pathway (Fig. 2.4). Receptor cytosolic domain is then activated by small GTPase ARF1 (cited above) to associate with adapter proteins on the cytosolic surface of *trans* Golgi membrane, recruiting clathrin for polymerization. Clathrin-coated vesicles also contain transmembrane proteins destined to the lysosome, which are sorted based on a tyrosine-containing motif. Vesicles pass by early and late endosomes, from where mannose-6-phosphate receptors are recycled to the *trans* Golgi, and cargo headed to lysosomes.

The endocytic pathway internalizes large molecules from the cell exterior or plasma membrane and controls the steady state of membrane components. There are two principal endocytic sorting hubs: the early and late endosomes. The early endosome is an anastomosing network of tubules that receives cargoes from the *trans* Golgi (secretory pathway) and the plasma membrane (endocytic pathway). In addition, it serves as the sorting station for cargoes that will be recycled backward to the plasma membrane, either directly or passing through the peri-Golgi recycling endosome. Some carrier vesicles may detach from early endosomes and progressively acquire membrane vesicles becoming a multivesicular body that will mature into late endosomes. Late endosomes as another important hub in the endocytic pathway can fuse with preexisting lysosomes. In certain cases, depending on the cell functional specialization, multivesicular bodies may fuse with the plasma membrane, releasing a collection of vesicles in the extracellular space (Fig. 2.4). These extracellular vesicles, collectively named exosomes, are particularly small (40–100 nm) and have been associated with different regulatory activities, functioning in a secretion-based manner of intercellular communication. Along the endocytic pathway, the identity of each compartment depends on several aspects. Importantly, the compartment lumen pH is a critically differentiating aspect as it relates to specific hydrolase activities. In addition, the small GTPase subfamily rab has several members that distinctively though not exclusively associate with the cytosolic face of compartment membrane.

Low-density lipoprotein receptor (LDLR) and transferrin receptor are prototypical receptors endocytosed by clathrin-coated vesicles that then fuse with early endosomes. *LDLR* loss-of-function mutations may cause familial hypercholesterolemia. Proprotein convertase subtilisin/kexin type 9 (*PCSK9*) binds to LDLR on the surface of liver cells, targeting the receptor to the lysosomes for degradation after endocytosis, reducing LDL-cholesterol uptake from plasma. Gain-of-function mutations in the *PCSK9* gene are associated with hypercholesterolemia, whereas loss of function markedly reduces plasma LDL-cholesterol, as LDLR recycling to the plasma membrane is increased. Likewise, *PCSK9* blockage by monoclonal antibodies increases LDL-cholesterol uptake from plasma (see Genetic case 5D).

---

*Genetic case 5D:* A man died of massive heart infarct at age 36. His son, now 28 years old, has arterial hypertension and hypercholesterolemia and is in use of statins and angiotensin-converting enzyme inhibitor. He heard on the news about novel treatments for patients with familial hypercholesterolemia. On a cardiology consult, he wants to know if his familial mutation could be therapeutically targeted.

---

Cargoes may differentially undergo endocytosis initiated in caveola microdomains and carried out in clathrin-independent vesicles (Fig. 2.4). During vesicle formation, caveolae proteins caveolin and cavin coat the membrane in an assembly believed to acquire a polyhedron structure. Caveolae vesicles may fuse in a caveosome or transiently with early endosomes. Caveolae also mediate transcellular shuttling of proteins between two different compartments, an important transport in endothelial cells. Caveolin dysfunction has been reported in patients with specific forms of lipodystrophy or muscular dystrophy.

## 2.6 Protein turnover and cell size control

---

Lysosomes are membrane-bound organelles containing enzymes that hydrolyze proteins, nucleic acids, lipids, and carbohydrates. Many lysosome enzymes are acid hydrolases functioning in acidic (5.0) pH maintained by proton pumps. Lack of activity in cytosolic neutral (7.2) pH controls undesired hydrolase activity outside the lysosome. Hydrolase substrates are the cell's own components as well as molecules endocytosed from outside the cell and those brought along the endocytic pathway such as plasma membrane components. In individuals afflicted by genetic diseases due to mutations in genes coding for lysosomal enzymes, collectively known as lysosomal storage diseases, undigested material accumulates in lysosomes (see Genetic case 6A).

---

*Genetic case 6A:* A 12-year-old male teenager from a high consanguinity family from a developing country is referred to the university hospital of the capital. He presents with slow-progression hepatosplenomegaly, lower limb pain, anemia, and thrombocytopenia. Different infectious diseases and hematological malignancies were ruled out. Bone MRI suggested necrotic lesions. Of note, patient's maternal and paternal grandmothers developed Parkinson's disease in their fifth decade, and present cognitive decline. An enzymatic assay disclosed low urine levels of glucocerebrosidase, confirming Gaucher disease, which is a lysosomal storage disease with autosomal recessive inheritance. The same mutation in the *GBA* gene, but in heterozygosity, may lead to Parkinson's disease.

---

Entire organelles may be targeted for degradation inside the cell in a process of turnover called autophagy. ER membrane partial fragmentation produces the isolation membrane or phagophore, a double membrane that grows in the cytosol, sequesters an organelle (e.g., mitochondrion in mitophagy), and seals itself as autophagosomes (Fig. 2.4). Autophagosomes fuse with lysosomes forming autophagolysosomes. Autophagy is a regulated pathway that provides energy for the starving cell. mTOR, a major activator of protein synthesis (Section 2.4), is also an inhibitor of autophagy. Autophagy proteins, normally termed following the acronym ATG, are organized in modules with sequential signaling and organization. In this sequence, upstream kinases regulate the first module, containing Unc5-like kinase 1 (ULK1), a nutrient sensor of the activity balance between mTOR and AMP-activated protein kinase. ATG modules regulate vesicle nucleation from the ER, elongation of the autophagosomal membrane, and autophagosome completion.

As mentioned (Section 2.5), UPR may retro-translocate unfolded proteins for degradation in the cytosol. Likewise, proteins synthesized by free ribosomes and delivered to the cytosol may be incorrectly folded, and marked for destruction. Besides the decoding center verification of the base pairing between mRNA codon and



tRNA anticodon, ribosomes have no other proofreading system to review if the amino acid used for peptide bond formation was indeed the correct one. In addition, although amino acyl-tRNA-synthases may verify and edit their own work of tRNA acylation, acylation errors may occur at 1:10,000 events. Consequently, proteins may be produced with amino acid substitutions without however any missense mutation in the corresponding gene. Based on these evidences, it is not surprising to observe that there are different proteins in the cell that act as chaperones assisting protein folding in the cytosol or the ER. Nevertheless, proteins may still remain unfolded or misfolded. Unfolded protein is directed toward the proteasome system for degradation, an abundant ATP-dependent protease dispersed in multiple copies throughout the cytosol.

The proteasome recognizes proteins covalently bound to poly-ubiquitins. Mono-ubiquitin is a 76-amino acid long peptide, and is firstly activated by covalent attachment to an ubiquitin-activating enzyme (E1). Ubiquitin is then transferred from E1 to ubiquitin-conjugating enzyme (E2) which associates with ubiquitin ligase (E3). As ubiquitin ligase will directly associate with the target protein, it needs to have some binding specificity. Therefore E3 constitutes a very large protein family encoded by more than 600 genes in the human genome. Mutations in E2 and E3 genes have been described in different clinical conditions. In particular, juvenile Parkinson's disease is a phenotype associated with unfolded protein aggregation (see Genetic case 6B). The proteasome is a large structure unbound by membrane, composed of a central, 20S cylinder containing the active site, and two external caps, 19S each. The 19S proteasome utilizes energy from ATP hydrolysis to unfold poly-ubiquitinated proteins and feed them to the narrow channel of the 20S proteasome. It keeps the entire substrate associated with it until fragmented in short peptides. The proteasome also digests proteins with regulated half-lives. One example of regulated poly-ubiquitination is considered below in the control of the cell cycle.

---

*Genetic case 6B:* A 28-year-old woman presents with posterior rigidity, postural instability, bradykinesia, and jaw and hand tremor. Her parents are deceased due to natural causes. She seeks treatment as she has been diagnosed with early-onset Parkinson's disease and shows concerns, as she has no family members who could help with daily duties. Juvenile Parkinson's disease has autosomal recessive inheritance and may be due to mutations in the ubiquitin ligase *PARK1* or *PARK2* genes, among other genes.

---

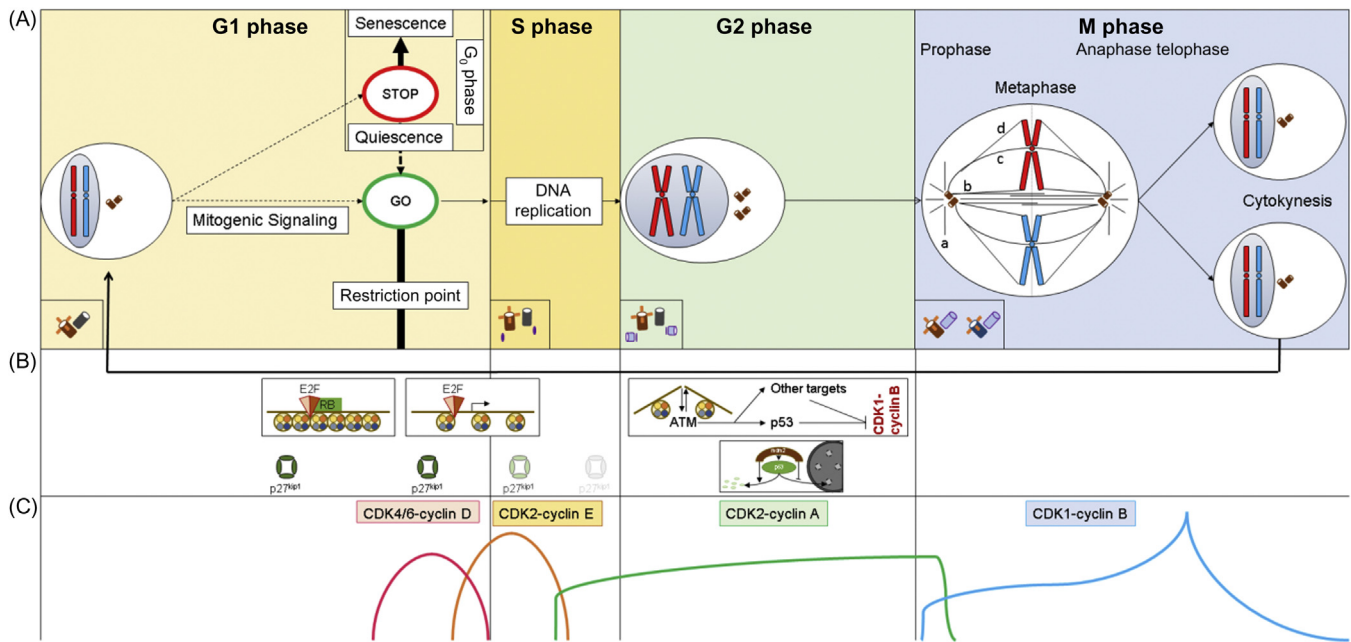
## 2.7 Microtubule-organizing centers and the cell cycle

MTOC are cell organelles having the ability to nucleate microtubules during interphase and mitosis, thus helping to establish cell architecture and polarization and, specifically, the mitotic spindle, cilium and flagellum. Centrosomes, the dominant and mostly recognized MTOC, are the major site of microtubule nucleation and anchoring in animal cells. Centrosomes are organelles not bound by membrane, organized around a pair of cylindrical centrioles, each composed of nine short microtubule triplets. Centrioles are important for maintenance of centrosome number, which may undergo amplification in a single cell under adverse conditions as malignant transformation. Centrosome amplification appears to affect the cell cycle and contributes to different aspects of tumor cells, such as centrosome enlarged size and abnormal shape, chromosomal instability, aneuploidy, disrupted cell division, and tissue invasion.

In centrosomes, a pericentriolar matrix surrounds centrioles linking them in orthogonal orientation (Fig. 2.5). In addition, the pericentriolar material contains several units of a ring-shaped complex centered at gamma-tubulin and associated proteins mimicking the format of the microtubule plus end (gamma-tubulin ring complex,  $\gamma$ TuRC), which is the basic piece responsible for nucleating microtubules. Spontaneous nucleation of microtubule is kinetically limiting, thus unfavorable. Hence,  $\gamma$ TuRC acts as a seed from which  $\alpha$ - and  $\beta$ -tubulin dimers polymerize in a polar manner, growing toward the microtubule plus end. On the opposite limit of the filament, microtubule minus ends are capped by  $\gamma$ TuRC, where fiber growth is limited. It has been demonstrated that  $\gamma$ TuRC recruitment to the pericentriolar matrix is necessary for the activation of its microtubule nucleation activity. In different species, cell organelles and structures other than the centrosome, such as the Golgi apparatus, the nuclear envelope and kinetochores, may contain proteins normally found in centrosomes displaying the ability to recruit  $\gamma$ TuRC and thus to function as MTOC.

### 2.7.1 The cell cycle

The cell cycle coordinates signaling events to duplicate the genetic material and divide the cellular contents into two identical daughter cells. After cell division, the cycling cell goes through G1 phase when it may grow



**FIGURE 2.5** The cell cycle. (A) Top panels depict cell cycle phases G1, G<sub>0</sub>, S, G2, and M. Entry in the cell cycle is indicated by the green sign “Go” in late G1 at the restriction point. A pair of homologous chromosomes is represented as red and blue diagrams. After S phase, in G2 sister chromatids in each chromosome represent duplicated DNA. In M phase a cell at metaphase is shown with spindle microtubules indicated as astral (a), polar (b), kinetochore (c), or chromosomal (d) microtubules. Duplicated centrioles are also indicated. An inset in the bottom left part of each panel highlights the centriole duplication cycle. (B) Mitogenic signaling needs to be sufficient to repress pRB and activate gene transcription by E2F TFs overcoming the restriction point. Degradation of the CDK inhibitor p27<sup>Kip1</sup> is represented in G1/S transition. In G2/M transition, DNA damage activates ATM kinase, which initiates DNA repair and signals to arrest the cell cycle by repressing CDK1-cyclin B through different protein intermediates, including p53. P53 activity is activated upon its release from MDM2 protein, allowing its nuclear import. Downregulated p53 reassociates with MDM2, which triggers its degradation. (C) Protein amount level variation for different cyclins and CDKs is indicated along the cell cycle and respective phases. CDK, Cyclin-dependent kinase.

and have its genome integrity verified. Environment and nuclear conditions permitting, the cell replicates DNA in S phase, moves to G2 phase for further growth and chromosome integrity check-up before entering M phase for novel cytokinesis (Fig. 2.5).

Centriole number is controlled during the cell cycle. At early G1 the centriole pair is composed of two mother (M) centrioles; one named M<sup>Old</sup> formed at least two cell cycles earlier, and the other one termed M<sup>New</sup> assembled in the previous cell cycle. Centriole duplication starts at G1/S phase transition, when each centriole initiates the formation of a procentriole through the nucleation of a ring of nine singlet microtubules on a disk-shaped structure containing a protein named centrin. The procentriole acquires complexity when matured into an array of nine triplet microtubules. The ninefold microtubule structure, then called daughter centriole, will lengthen during G2 phase and G2/M transition. M<sup>New</sup> centriole becomes M<sup>Old</sup> centriole at early mitosis when it acquires protein appendages, a characteristic of the oldest centriole in a pair. Each pair of centrioles, now harboring one mother centriole and a daughter centriole, migrate apart over the surface of the nuclear envelope, setting up the two poles of the mitotic spindle. After cytokinesis, each daughter centriole then called M<sup>New</sup> centriole pairs in orthogonal placement to a M<sup>Old</sup> centriole (Fig. 2.5).

Protein activity regulation by phosphorylation, degradation upon poly-ubiquitination, and specific protein–protein interactions controls and defines the unidirectional cell cycle. Of note, kinases dependent on small proteins named cyclins (cyclin-dependent kinases, CDKs) are key regulators responsible to overcome cell cycle suppression at each phase, depending on the rise of specific cyclin levels. Upon entering G1, cells are not prone to cycle unless sufficient mitogenic signaling allows them to overcome a restriction point at the end of this phase. At this point, cells assess environment conditions (internal and external stimuli) to commit to another cycle and enter S phase. On the other hand, cells that underwent differentiation in the previous cycle as well as those under unfavorable environment conditions such as lack of nutrients or mitogenic factors may leave the cycle at G1 before its restriction point. These cells enroll a nonproliferative phase named G<sub>0</sub>, which may last hours or days for quiescent cells or the lifetime of the senescent cell. Although not proliferative, G<sub>0</sub> cells are active and expend energy (Fig. 2.5).

A major control at G1 checkpoint is the repression by retinoblastoma tumor suppressor protein family, pRB, on the TF E2F family inhibiting the expression of genes required for cell cycle progression. In the lack of mitogens, pRB binds E2F, which associates with promoters of regulated genes, recruiting histone deacetylases thus compacting chromatin and repressing transcription. Signaling initiated by mitogens leads to the expression of specific cyclins and CDKs. Rising levels of cyclin D and CDK4/6 trigger phosphorylation of pRB by the latter, which is released from promoters occupied by E2F, bringing with it the histone deacetylation complex. Consequently, histone acetylation is activated, chromatin is decondensed, what up-regulates transcription of genes encoding proteins that function in the transition of G1 to S phase, such as cyclin E and CDK2. Cyclin E/CDK2-initiated transition from G1 to S phase activates p27<sup>kip1</sup> degradation, an inhibitor of cyclin D and cell cycle progression (Fig. 2.5).

Cell environmental genotoxic effects are most frequently repaired at G1 or G2 phases, respectively before replicating the genetic material (S phase) or segregating sister chromatids in mitosis. Moreover, stalled DNA replication forks are repaired in S phase. Repair takes place by activation of specific checkpoint-associated signaling cascades. At both G1/S and G2/M transitions, p53, a tumor suppressor protein, is a central regulator of genome stability. DNA damage activates ATM (ataxia telangiectasia mutated) and ATR (ataxia telangiectasia and rad3-related) kinases. ATM/ATR phosphorylates p53, releasing it from MDM2 protein. High levels of p53 activates transcription of the *CDKN1A* gene, which expresses the protein p21, a CDK inhibitor, maintaining pRB association with E2F thus blocking the transition from G1 to S phase. Possible outcomes for the cell cycle arrest are DNA repair, cell senescence or programmed cell death. At G2 checkpoint, activated p53 drives the expression of a number of genes, prolonging cell cycle arrest. ATM also utilizes other targets that culminate in cyclin B-cdk1 inhibition. Importantly, ATM interacts with different DNA-repair protein complexes, hence directly assisting the repair process. Inherited mutations in the *TP53* gene leads to Li–Fraumeni syndrome, a cancer predisposition condition in which the cell cycle arrest and DNA-repair response to DNA damage are impaired (see Genetic case 7A).

---

*Genetic case 7A:* A 29-year-old female presents with a ductal carcinoma in the left breast. Her father had an early diagnosed and treated osteosarcoma when he was a child and is otherwise healthy (age 50). Her father's sister died of brain cancer at age 44, and paternal grandfather had multiple tumors, including a colorectal adenocarcinoma, the cause of his death. A missense mutation was identified in the *TP53* gene, confirming the diagnosis of Li–Fraumeni syndrome. Her father also tested positive. Counseling was provided.

---

Mitosis (M phase) is the division of a somatic cell into two daughter cells. This phase can be subdivided into six phases: prophase, prometaphase, metaphase, anaphase, telophase, and cytokinesis. Dramatic changes taking place in the nucleus and cytoplasm during mitosis are regulated by kinases, remarkably the cyclin B-CDK1 complex among others. As seen, centrosomes define the two poles of the mitotic spindle in prophase. Centriole changes during the cell cycle have been explained above. Microtubule dynamics changes rapidly in M phase. In early prophase, microtubules preferentially disassemble and shrink and, in parallel, there is an increase in the number of new filaments emanating from the centrosome. Nuclear changes include chromosome condensation, envelope breakdown, and nucleolus disassembly. Prometaphase is characterized by spindle fibers penetrating the cell nucleus through the disintegrating nuclear envelope and attaching to chromosome kinetochores. At metaphase, when condensed chromosomes align on spindle equator, there are four types of stabilized microtubules: (1) kinetochore microtubules attached to the condensed chromosomes at the kinetochore; (2) chromosomal microtubules associated with chromosomal ends; (3) polar microtubules overlapping each other at the center of the cell; and (4) astral microtubules extending outward the centrosomes to the periphery of the cell displaying plus ends (Fig. 2.5). At anaphase, motor proteins associated with spindle microtubules mediate the pulling out of sister chromatids to opposite poles of the spindle. After reassembly of the cell nucleus envelope at telophase, spindle poles continue to separate due to the sliding between polar microtubules, and an actomyosin contractile ring is formed at the cleavage plane so that cytokinesis can take place.

### 2.7.2 Primary cilium

The primary cilium is a solitary organelle that protrudes on the surface of most metazoan cell types at senescence or quiescence (G<sub>0</sub>) or from G1 to early mitosis in cycling cells. The axoneme is the essential structure of

flagella and cilia, composed of a central pair of microtubules and nine outer microtubule doublets radially disposed, enclosed by the plasma membrane. Nonmotile cilium lacks central microtubules. The minus ends of cilium and flagellum microtubules are anchored in a basal body, which is equivalent to the centriole. In fact, the basal body is the mother centriole from a pair that has been modified to connect the plasma membrane at the basis of the primary cilium. This interaction is mediated by the centriole distal appendages or transition fibers. Due to the representation of the axoneme as the elongation of part of microtubules from the basal body, axoneme and basal body are commonly referred to as the functional unit of the primary cilium.

In epithelial cells the cilium functions as fluid sensors, chemosensors, as well as sensors of light, temperature, or osmolality. Dysfunctions of ciliary proteins leading to heterogeneous clinical presentations, such as deafness, cystic kidney diseases, cardiac malformation, or hydrocephalus, have been named altogether as ciliopathies (Table 2.1). Like some other renal cystic disorders, autosomal dominant polycystic kidney disease may be classified as ciliopathy due to loss-of-function mutations affecting specific ciliary proteins, polycystins 1 and 2 (see Genetic case 7B). Polycystin 2 is a ciliary calcium channel that responds upon mechanosensation, such as cilium bending, due to fluid flow in the nephron lumen. Calcium in the cilium acts as a second messenger initiating intracellular signaling. In addition, as chemosensors, cilia may play signaling roles originated by trophic factors.

---

*Genetic case 7B:* A 45-year-old man has episodic lower back pain. Ultrasonographic studies revealed few cysts in liver and pancreas, and multiple cysts bilaterally in the kidneys. Glomerular filtration rate was at 50 mL/min, and there was microscopic hematuria with no proteinuria. He had recently initiated treatment for arterial hypertension. Autosomal dominant polycystic kidney disease usually fully manifests in the fourth or fifth decade and is due to mutations in the *PKD1* or *PKD2* genes.

---

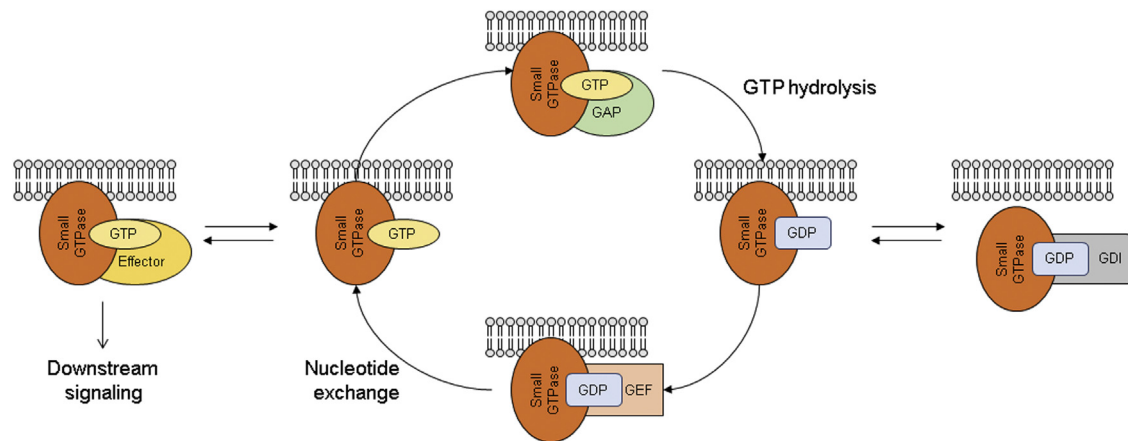
## 2.8 Energy production and oxidative stress

Biological processes require energy. Many chemical reactions that take place in the cell are energetically unfavorable. The nucleotide ATP plays a central role as storage of free energy in the cell. The hydrolysis of ATP into ADP plus inorganic phosphate yields free energy from the phosphate high-energy bond. In addition to ATP, GTP is a reliable storage of energy. In different cell processes the hydrolysis of GTP into GDP and phosphate is performed by small GTPases (Table 2.2). GTP-bound small GTPase may interact with specific protein effectors, activating a biological process. Alternatively, a GTPase-activating (GAP) domain-containing protein may associate with GTP-bound small GTPase, catalyze GTP hydrolysis, thus inactivating the small GTPase and downregulating its signaling cascade. GDP must be exchanged by GTP to restore the small GTPase activity, what is achieved by a member of the GEF family. Under cell stress conditions, GDP-bound small GTPase may be trapped by a specific GDP dissociation inhibitor protein, which keeps the GTPase inactivated until cell homeostatic conditions return (Fig. 2.6). Some proteins from the Ras monomeric small GTPase protein superfamily are listed on Table 2.2 as well as the biological process regulated when associated with GTP. The large GTPase protein family comprehends other classes of GTPase proteins, such as the guanine nucleotide-binding (G) protein subunit alpha from heterotrimeric G proteins, involved in signal transduction (Section 2.1); as well as the dynamin GTPase superfamily that regulates membrane fusion and fission (Section 2.5); and the GTPase translation factor family, fundamental in protein synthesis (Fig. 2.3; Table 2.2, Section 2.6). Different from the small GTPase superfamily, some abovementioned GTPases do not depend on GAP and GEF proteins. Some of them display lower affinity to GTP or GDP and have internal mechanisms to stabilize the transition state during GTP hydrolysis. Others as the G proteins have a specific class of effector and GAP proteins. In conclusion, purine triphosphate is an efficient form of energy storage and availability, and is associated with proteins with intrinsic GTPase activity, thus rapidly providing regulation of biological processes.

### 2.8.1 Mitochondria

The mitochondrion is an organelle bound by two membranes that define an intermembranous space and a matrix, which contains 55S ribosomes and several copies of a circular DNA molecule that constitutes the mitochondrial genome. Mitochondrial DNA carries intronless genes for mitochondria-owned rRNA and tRNA sets, as well as for few mitochondrial membrane proteins. Mitochondrial genes are disposed in economical





**FIGURE 2.6** The small GTPase protein cycle. Small GTPases are peripheral membrane proteins associated with the cytoplasmic face of the plasma membrane or organelle membranes through posttranslational lipid modification. They cycle between GTP- or GDP-bound states. Binding to GTP increases its affinity to specific effector proteins mediating downstream signaling or to GAP catalyzing GTP hydrolysis, inhibiting the GTPase. Under cell stress conditions, GDP-bound small GTPases may be temporarily captured to an inactivated state by association with a protein that functions as a GDI. Homeostatic conditions unbinds GDI, and upstream signaling drives GDP-associated small GTPase to interact with a specific GDP-GTP exchange factor (GEF), which exchanges GDP to GTP, restoring the active form of the GTPase. GTP-bound small GTPase reinitiates the cycle by effector or GAP association. *GAP*, GTPase-activating proteins; *GDI*, GDP dissociation inhibitor; *GTP*, guanosine 5'-triphosphate; *GDP*, guanosine 5'-diphosphate.

orientation in the circular DNA molecule, if compared to the eukaryotic nuclear genome organization. In fact, common aspects between mitochondria and prokaryotic cells, such as the compact genome, ribosome structure, and gene expression mechanisms, corroborate the endosymbiosis model that predicts that eukaryotic mitochondria have originated from an aerobic bacterium that had been endocytosed by and symbiotically coexisted in a eukaryotic, anaerobic, and unicellular ancestor. The human mitochondrial genome is 16,569-bp long and codes for 13 proteins. Among the approximate 21,000 protein-coding genes from the human nuclear genome, nearly 1000 encode proteins that are synthesized in the cytoplasm and imported into the mitochondria (Human Genome Annotation Release 108, accessed in December 2017). Nuclear-encoded proteins targeted to mitochondria have a distinct sorting system independent on the secretory pathway, not presented in this chapter.

Glycolysis and fatty acid oxidation feed high-energy chemical intermediates to the citric acid cycle that takes place in the mitochondria matrix. Those metabolic pathways basically converge to produce acetyl-CoA, which enters the citric acid cycle. Each acetyl-CoA molecule entering the cycle produces three molecules of NADH (nicotinamide adenine dinucleotide), one molecule of FADH<sub>2</sub> (flavin adenine dinucleotide), and two molecules of carbon dioxide (CO<sub>2</sub>). NADH and FADH<sub>2</sub> electrons pass through the electron transport pathway, a series of proteins in the inner membrane of mitochondria, finally leading to the reduction of molecular oxygen to water by four electrons. The energy released along the electron transport pathway is stored in the form of a proton gradient generating an electrical potential across the inner membrane. ATP synthase, a transmembrane protein complex associated with the electron transport pathway in the mitochondrion inner membrane, uses the electrochemical proton gradient to drive ATP production in the mitochondrion matrix upon pumping hydrogen, ADP and phosphate from the intermembranous space to this compartment. Altogether, the final reactions of oxygen reduction and ADP phosphorylation are known as oxidative phosphorylation. Increasing energetic demand by the cell is met by higher rates of mitochondria biogenesis and fusion, whereas reduced energy demand relates to the removal of mitochondria by fission or autophagy (mitophagy).

A by-product of the mitochondrial electron transport is reactive oxygen species (ROS), which at low levels may function in cell signaling. Mitochondrial antioxidant systems prevent DNA, lipid, and protein damage by high levels of ROS. Superoxide (O<sub>2</sub><sup>•−</sup>) is the most frequent ROS and is produced by molecular oxygen reduction by a single electron. One important antioxidant defense is the activity of the superoxide dismutase (SOD). The *SOD1* gene encoding this enzyme is mutated in a subset of patients with familial form of amyotrophic lateral sclerosis (see Genetic case 8A). Increased ROS levels can have impact on mitochondrion functional activities, decreasing ATP synthesis.

---

*Genetic case 8A:* A 48-year-old man reports trouble walking and writing. Upon examining, he presents bilateral wasting of the shoulder girdles, atrophic weakness of hands, lower limb hyperreflexia and stiffness, tongue fasciculation, and no sensorial changes. Familial amyotrophic lateral sclerosis (onset around 50 years) corresponds to up to 10% of all amyotrophic lateral sclerosis cases and may have autosomal dominant or recessive inheritance. Approximately 20% of familial cases present mutation in the *SOD1* gene

---

Mitochondria play important roles in apoptosis regulation especially as homeostasis sensor triggering the intrinsic pathway of cell death. Proapoptotic stimuli, for example, increase in ROS levels or in AMP/ATP ratio, activate the insertion of proapoptotic Bcl-2 family members Bax and Bak in mitochondrial outer membrane. Their oligomerization causes the release of proapoptotic factors, such as cytochrome C, from mitochondria intermembrane space into the cytosol. Cytochrome C triggers the formation of the apoptosome, activating downstream caspases. Apoptotic cell death characteristics include chromatin condensation, cell shrinkage and fragmentation. Apoptosis intrinsic and extrinsic pathways will not be presented in this chapter due to its complexity and importance to themes covered in other chapters of this book.

### 2.8.2 Peroxisomes

Peroxisomes are dynamic organelles involved by a single membrane that limits a lumen containing several oxidative enzymes. Peroxisomes may be involved in different anabolic and catabolic processes, such as biosynthesis of ether phospholipids and bile acids,  $\alpha$ - and  $\beta$ -oxidation of fatty acids, and detoxification of glyoxylate and ROS. The organelle is named after catalase catalysis of hydrogen peroxide decomposition, a frequent peroxisome detoxification reaction. Peroxisomal proteins are synthesized by free cytoplasmic ribosomes, and posttranslationally targeted to the preperoxisome, a vesicle budded from the ER carrying import receptors, named PEX. Alternatively, preexisting peroxisomes can grow by importing proteins and lipids or may divide by GTPase-dependent fission (Table 2.2). Mutations in genes encoding at least 13 different PEX proteins responsible for peroxisome biogenesis may cause Zellweger spectrum disorders, with severity varying from mild with adolescence/adult onset to severe in the neonate (see Genetic case 8B). Patients accumulate very long-chain fatty acids and bile intermediates in the plasma.

---

*Genetic case 8B:* A low-birth-weight newborn presents with severe hypotonia, suckling difficulty, large fontanels, seizures, jaundice, hepatosplenomegaly, and bloody stools. Brain MRI additionally identified perisylvian polymicrogyria. Laboratory tests for the assessment of peroxisomal biogenesis disorders included plasma levels of very long fatty acids (C22:0, C24:0, and C26:0) and next-generation sequencing of targeted gene panel for inborn errors of metabolism.

---

## 2.9 Summary

---

DNA is a unique molecule that stores genetic information for cell growth, proliferation and function. The cell machinery has the ability to make an accurate copy of DNA and transmit it to daughter cells following a unidirectional cycle that may undergo regulation according to environment conditions. Changing environment conditions allow for cell adaptation based on signaling systems that rely on developmentally regulated genetic programs. Additional layers of molecular surveillance have evolved as major mechanisms maintaining the species genome integrity, destroying nonsense mRNA or misfolded proteins, and controlling the cell size. In cell homeostasis, protein GTPases combine energy expenditure to organization of key biological processes, regulating gene expression, molecule trafficking, cytoskeleton assembly and disassembly, and membrane fusion and fission, among other functions.

Our molecular understanding of the cell has been accelerated after the finishing of the human genome sequencing. This breakthrough in addition to the development of next-generation DNA sequencing platforms, in use for the last decade, has catalyzed the discovery of novel genetic causes of human disease, and has allowed for different experimental approaches employing genetically modified cells and model organisms. Moreover, the well-conserved structure of the eukaryotic cell across four biological kingdoms and the availability of genome sequences from several species have made possible vast comparative morphological and gene expression

analyses among different tissues and organisms. Therefore one can view human pathologies as dysfunctions of a cell organelle, as exemplified in this chapter, or a molecular pathway. On the one hand, this knowledge may corroborate and strengthen the molecular pathogenesis of a group of diseases with similar clinical classification. On the other hand, pathologies previously viewed as unrelated clinical entities may share a common organelle dysfunction allowing for novel avenues for therapeutic testing.

## Bibliography

- Alberts et al., 2014 [Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, et al. Molecular biology of the cell. 6th ed. New York: Garland Science; 2014.](#)
- Aviram and Schuldiner, 2017 [Aviram N, Schuldiner M. Targeting and translocation of proteins to the endoplasmic reticulum at a glance. J Cell Sci 2017;130:4079–85.](#)
- Cherfils and Zeghouf, 2013 [Cherfils J, Zeghouf M. Regulation of small GTPases by GEFs, GAPs and GDIs. Physiol Rev 2013;93:269–309.](#)
- Cooper and Hausman, 2015 [Cooper GM, Hausman RE. The cell: a molecular approach. 7th ed. Oxford: Sinauer Associates \(Oxford University Press\); 2015.](#)
- Dekker and Misteli, 2015 [Dekker J, Misteli T. Long-range chromatin interactions. Cold Spring Harb Perspect Biol 2015;7\(10\):a019356.](#)
- De-Silva et al., 2015 [De-Silva D, Tu Y-T, Amunts A, Fontanesi F, Barrientos A. Mitochondrial ribosome assembly in health and disease. Cell Cycle 2015;14:2226–50.](#)
- Elvekrog and Walter, 2015 [Elvekrog MM, Walter P. Dynamics of co-translational protein targeting. Curr Opin Chem Biol 2015;29:79–86.](#)
- Ferguson and De Camilli, 2012 [Ferguson SM, De Camilli P. Dynamin, a membrane-remodelling GTPase. Nat Rev Mol Cell Biol 2012;13\(2\):75–88.](#)
- Fujiki et al., 2014 [Fujiki Y, Okumoto K, Mukai S, Honsho M, Tamura S. Peroxisome biogenesis in mammalian cells. Front Physiol 2014;5:307.](#)
- Herst et al., 2017 [Herst PM, Rowe MR, Carson GM, Berridge MV. Functional mitochondria in health and disease. Front Endocrinol 2017;8:296.](#)
- Hetz and Saxena, 2017 [Hetz C, Saxena S. ER stress and the unfolded protein response in neurodegeneration. Nat Rev Neurol 2017;13\(8\):477–91.](#)
- Izumi et al., 2016 [Izumi K, Brett M, Nishi E, Drunat S, Tan E-S, et al. \*ARCNI\* mutations cause a recognizable craniofacial syndrome due to COPI-mediated transport defects. Am J Hum Genet 2016;99:451–6.](#)
- Klower et al., 2015 [Klower FCC, Berendse K, Ferdinandusse S, Wanders RJA, Engelen M, Poll-Tee BT. Zellweger spectrum disorders: overview and management approach. Orphanet J Rare Dis 2015;10:151–62.](#)
- Kressler et al., 2017 [Kressler D, Hurt E, Bassler J. A puzzle of life: crafting ribosome subunits. Trends Biochem Sci 2017;42\(8\):640–54.](#)
- Laird et al., 2017 [Laird DW, Naus CC, Lampe PD. Snapshot: connexins and disease. Cell 2017;170:1260–1.](#)
- Lamaze et al., 2017 [Lamaze C, Tardif N, Dewulf M, Vassilopoulos S, Blouin CM. The caveolae dress code: structure and signaling. Curr Opin Cell Biol 2017;47:117–25.](#)
- Maib et al., 2015 [Maib H, Smythe E, Ayscough K. Forty years on: clathrin-coated pits continue to fascinate. Mol Biol Cell 2015;28:843–7.](#)
- Pollard et al., 2016 [Pollard TM, Earnshaw WC, Lippincott-Schwartz J, Johnson G. Cell biology. 3rd ed. Amsterdam: Elsevier; 2016.](#)
- Reiter and Leroux, 2017 [Reiter JF, Leroux MR. Genes and molecular pathways underpinning ciliopathies. Nat Rev Mol Cell Biol 2017;18:533–47.](#)
- Rinn and Guttman, 2014 [Rinn J, Guttman S. RNA and dynamic nuclear organization. Science 2014;345:1240–1.](#)
- Soutorina, 2018 [Soutorina J. Transcription regulation by the mediator complex. Nat Rev Mol Cell Biol 2018;19\(4\):262–74.](#)





# Molecular basis of clinical metabolomics

*Diego F. Gomez-Casati and María V. Busi*

Centre for Photosynthetic and Biochemical Studies (CEFOBI-CONICET), National University of Rosario, Argentina

## 3.1 Introduction

Metabolomics, the comprehensive high-throughput analysis in which all the metabolites in a biological system are identified and quantified [1,2], has emerged as a functional genomics methodology which helps to better understand the complex molecular interactions in biological systems [3]. Metabolomics strategies are closely related to other omics, such as genomics, transcriptomics, and proteomics, because it represents the logical progression from large-scale analysis of nucleic acids and proteins at the systems level [4,5]. Recently, metabolomics became essential as a tool in medicine, with an important role in the diagnosis and prevention of numerous diseases and in the design of strategies to carry out an adequate treatment for each pathological condition [6].

Metabolomics allows us to evaluate the physiological state of a cell in the context of its environment and considering the modification of enzyme kinetics and changes in metabolic reaction rates [7–9]. Thus compared with genomics or proteomics, metabolomics reflects changes in phenotype and therefore function. The “omic” sciences are, however, complementary as “upstream” changes in genes and proteins are measured “downstream” as changes in cellular metabolism [7,10,11]. Other features of metabolomics are similar to those of proteomics and transcriptomics, including the ability to assay biofluids or tumor samples and the relatively inexpensive, rapid, and automated techniques once start-up costs are taken into account.

Metabolites are the end products of the different cellular processes. Metabolite levels are related to the response of biological systems to changes in genetic and/or physiological conditions or environmental changes. On the other hand, the metabolome refers to the complete set of metabolites that are present in a biological system which participates in biochemical pathways required for its normal function [1,12].

Thus metabolomics is considered the ultimate level of post-genomic analysis because it can reveal changes in metabolite levels that are controlled by only minor changes in gene expression measured using transcriptomics techniques and/or by analyzing the proteome that could reveal posttranslational control mechanisms on enzyme activity [13]. In the last years, metabolomics has not only been used for medical applications but also an extensive research on metabolomics was carried out in many organisms, such as bacteria, fungi, parasites, and plants, because they produce a vast chemical diversity of compounds that may be beneficial for humans, including food, pharmaceuticals, and industrial raw materials [14].

This chapter describes different applications of metabolomics in clinical and medical practice, such as the identification and use of markers related to genetic inborn errors, the use of markers for the diagnosis, prevention, prognosis, and treatment of different human diseases, such as cancer, heart diseases, and neurological pathologies, and for the study and diagnosis of mitochondrial diseases.

## 3.2 Clinical applications of metabolomics

### 3.2.1 Inborn errors of metabolism

Inborn errors of metabolism are a group of more than 500 rare genetic diseases caused mainly by defects in many metabolic pathways [15]. The variety of metabolic pathways that may be affected determines how difficult

the diagnosis of this group of diseases is. These disorders comprise defects or deficiency of the catalytic activity of enzymes and deficiency in cofactors which leads to the abnormal accumulation of metabolites. The alteration of metabolite levels (accumulation or depletion) could produce a toxic effect on the organism and could also lead to the production of an adaptive and sometimes unwanted response. In addition, it is possible that the patient shows some symptoms and signs due to this alteration [16].

In general, newborns may not present some of the symptoms caused by a metabolic deficit because the metabolites can be translocated to the maternal placenta during gestation and thus be excreted; however, depending on the degree of gravity of the disease, the symptoms will appear within a few days of birth [16].

Some of the disorders comprise amino acid deficiencies, organic acid, and fatty acid defects. However, currently, the metabolic inborn error diseases are classified according to which organ is affected (i.e., cardiac, neurological, hepatic disease) or to which organelle is affected (i.e., mitochondrial and peroxisomal disorders) [17]. In the recent years the combination of genomic, transcriptomic, and metabolomic studies leads to significant progress in the identification, evaluation, and treatment of these group of diseases [18].

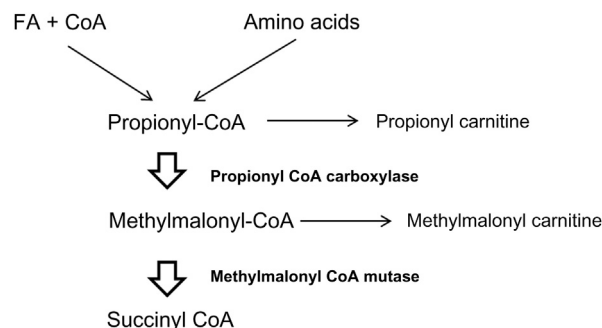
Within the most studied pathologies is the group of organic acid deficiencies that lead to the accumulation of mono-, di-, or tricarboxylic acids. This group of pathologies includes methylmalonic acidemia (MMA), propionic and isovaleric acidemia, maple syrup urine disease, 3-hydroxymethyl glutaryl-CoA lyase (HMGCL) deficiency, beta-ketothiolase deficiency, and glutaric aciduria (GA) [17].

The MMA is an autosomal recessive inborn error of metabolism caused by the deficiency of methylmalonyl-CoA mutase or by the deficiency of its cofactor 5'-deoxyadenosylcobalamin (Fig. 3.1) [19]. This group of patients has early mortality and shows multisystemic manifestations, such as pancreatitis, renal and liver insufficiency, and hepatomegaly [19].

The early and rapid diagnosis of these pathologies to carry out an effective treatment is extremely important. In most cases, neither diet, vitamin supplementation nor surgery treatment, such as liver transplantation, can prevent the multiorgan defects. The diagnosis of integrating proteomic and metabolomic studies has facilitated the rapid and economical evaluation of a broad spectrum of metabolites in various body fluids. Particularly, in MMA by using gas chromatography coupled to mass spectrometry (GC-MS), the levels of methylmalonic acid and methylcitric acid were increased [17]. Furthermore, using a proteomic approach, a downregulation of several proteins related to energetic metabolism, such as cytochrome *c* and succinyl CoA ligase, has been found, whereas other proteins involved in oxidative stress responses, such as superoxide dismutase, as well as the production of ROS were found to be increased, indicating that there is an impairment of cellular energy metabolism [20].

Propionic acidemia (PA) is a disease caused by a defect in propionyl-CoA carboxylase, which catalyzes the conversion of propionyl-CoA to methylmalonyl-CoA in the presence of biotin as a cofactor (Fig. 3.1) [21]. This and the MMA were the first two diseases that have been diagnosed using MS. An increase in propionyl carnitine was observed in PA and this metabolite is currently used as the primary biomarker for this disease [22]. Other metabolites, such as unsaturated acylcarnitines, isovaleryl carnitine, and  $\gamma$ -butyrobetaine, were associated with MMA and PA [22].

Maple syrup disease (MSD) is another metabolic disorder which causes an increment in the levels of some branched amino acids, such as leucine, isoleucine, and valine (and their toxic ketoacid products), due to the cellular inability to metabolize them. The disease is caused by a deficiency of the branched-chain ketoacid dehydrogenase complex and characterized by the presence of a smell similar to maple syrup in the urine of infants and brain dysfunction [23]. There are different types of MSD disease, classic, intermediate, intermittent, and a type of



**FIGURE 3.1** Pathway of production of succinyl CoA from propionyl CoA. FA, Fatty acids.

disease that responds to thiamine treatment, related to the levels of residual enzyme activity. The disease could be treated with adequate diet; however, depending on the severity of the disorder, patients could have different episodes with stress and infection [24].

Recently, an untargeted metabolic analysis for the screening of inborn metabolic errors reported the identification of several metabolites that could be used as markers for MSD and other disorders [25]. Thus increased levels of allo-isoleucine, 4-methyl-2-oxopentanoate, 3-methyl-2-oxovalerate, leucine, 2-hydroxy-3-methylvalerate, isoleucine, and alpha-hydroxyisovalerate, whereas a decrease of 3-hydroxyisobutyrate, isovalerylcarnitine, 2-methylbutyrylcarnitine, and hydroxyisovalerylcarnitine, were observed in MSD disease [25].

The HMGCL deficiency is a disease characterized by the inability to metabolize leucine and also affect ketogenesis. The disorder is caused by mutations in the *HL* gene. Patients with a deficiency of this enzyme have a reduced capacity to synthesize ketone bodies and the symptoms are related to the presence of low blood glucose levels causing metabolic acidosis and weak muscle tone [26]. LC–MS analysis revealed different metabolites that were found increased in patients with HMGCL deficiency, such as 3-methylglutaryl carnitine, hydroxyisovalerylcarnitine, glutaryl carnitine, beta-hydroxyisovalerate, and isovalerylcarnitine [25]. If this metabolic disorder is not treated early, it can cause seizures, breathing problems, coma, and death.

Another common metabolic disease classified as an inborn error of metabolism is the GA. This is an inherited disorder caused by deficiency of glutaryl-CoA dehydrogenase which is involved in the catabolism of L-lysine, L-hydroxylysine, and L-tryptophan [27]. GA disease is characterized by brain damage and macrocephaly and the first symptoms occur early in the infancy. Metabolites found to be increased in GA comprise glutaric acid, 3-hydroxyglutaric acid, glutaconic acid, and glutaryl carnitine, which can be detected by GC–MS or tandem MS [27,28].

In many cases, more than one metabolic error may occur. Recently, a case was reported in a newborn patient who showed some symptoms due to two disorders that were jointly presented: GA and isobutyrylglycinuria (IBD) [28]. IBD is caused by the deficiency of isobutyryl coenzyme dehydrogenase, an enzyme involved in valine catabolism. Using tandem-MS analysis, Popek et al. reported the accumulation of acylcarnitines C5DC and C4, 3-hydroxyglutaric acid, and isobutyrylglycine in body fluids. Thus the use of acylcarnitine was proposed as a biomarker for the diagnosis of IBD [28].

### 3.2.2 Metabolomics in cancer and other human diseases

Recent genomic studies have identified many genetic polymorphisms that convey an increased risk for developing diabetes, coronary artery disease, rheumatoid arthritis, cancer, and other common diseases [29–32]. However, little can be inferred about the mechanisms that cause these diseases only by the association of genotypes with clinical outcomes. Therefore it is crucial to screen large populations of patients in order to obtain sufficient statistical power for the identification of new disease-causing genetic variants, as recent genome-wide association studies with up to 18,000 patients have demonstrated [33–35]. In this sense, metabolomic studies may contribute to solving these problems [33].

Biochemical measurements of particular intermediate phenotypes on a continuous scale can be expected to provide more details on potentially affected pathways and to be more directly related to the etiology of the disease. It thereby provides a functional analysis of the physiological state of the human body. Genetic variants that associate with changes in the homeostasis of molecules, such as lipids, carbohydrates, or amino acids, are not only expected to display much larger effect sizes due to their direct involvement in metabolite conversion modification but may also provide access to the underlying molecular disease-causing mechanisms [33].

The recent interest of metabolomics in oncological studies initially stemmed from the claim in the late 1980s that many cancer metabolites could be identified by nuclear magnetic resonance (NMR) spectra of blood samples [36]. Recently, the progress in many NMR spectroscopy and mass spectrometry (MS) techniques, the most accepted methods used for the quantification of metabolites, has improved the sensitivity and resolution of analytic assays on patients samples in attempts to achieve a comprehensive biochemical assessment. Because cancer cells show a highly unique metabolic phenotype, the development of specific biomarkers is possible that could be used in identifying fingerprints and metabolic profiles to detect the presence of tumors, select the appropriate treatments, and/or assess the pharmacodynamic effects of therapy [7,11,37–42].

Probably, the best application of metabolomics in cancer diagnostics is breast cancer. Biomarkers are widely used in clinical medicine for prognostic or predictive purposes. Two examples of this are the estrogen receptor and the human epidermal growth factor receptor 2 (*Her2/neu*) [11]. In general the metabolic biomarkers are first

explored preclinically using animal and human cell cultures, followed by validation in fluids or tumoral tissues. Thus in the last few years, the metabolome of different tumors began to be characterized with more detail. Using standard metabolomic methods, tumors, in general, display elevated phospholipid levels [characterized by an elevation of total choline-containing compounds (tCho) and phosphocholine] and increased glycolytic capacity, including increased utilization of glucose carbons to drive synthetic processes, high glutaminolytic function, and overexpression of a glycolytic isoenzyme, the pyruvate kinase type M2 (M2-PK; [38,43,44]). The M2-PK enzyme may be of particular interest because its inactive dimeric form is dominant in tumors and thus this type has been named M2-PK tumors. Interestingly, lipid metabolic profiles have been documented to be 83% accurate at discriminating between cancer patients and controls, using NMR-based metabolomics of blood samples [45]. Importantly, in vivo, tCho determination via MRSI has detected breast, prostate, and brain tumors and correlates well with diagnosis via dynamic contrast enhanced-MRI [11,46–49].

Metabolomics was also used for the identification and diagnosis of colorectal cancer (CRC). CRC is one of the most common cancers worldwide among humans. The progress of this disease is generally slow, with few apparent symptoms; however, the detection is usually done in advanced stages [50]. The colonoscopy is important for early detection of CRC but it has the disadvantage that it is an invasive technique. A promising approach is to detect biomarkers in urine, plasma, or fecal samples to improve the early diagnosis of the disease [50].

At present, whereas the incidence of patients with lung cancer is decreasing, it is still cancer with the highest mortality rate. Survival of patients is low (less than 15% at the end of 5 years) [51], because the CRC does not present clear early symptoms and the pathology is quite complex, since there are different types and subtypes of cells that cause the disease. Thus one of the objectives of research in lung cancer is the identification of biological markers that allow an early and precise diagnosis of the disease in order to carry out an appropriate treatment. There are numerous works on the search for these biomarkers. Some of the metabolites that have been found differentially increased in patients with lung cancer include glycerol, phosphoric acid, hydroxylamine, threonine, phosphatidyl ethanolamine, putrescine, choline, acetone, lactic acid, inositol, valine, diethyl spermine, and xylose, among others (reviewed in [51]).

Despite the advances made in the area of cancer metabolomics, there is much information that remains to be elucidated. For example, among the different types of tumors, there are variations in the metabolic profiles of different compounds, such as alanine, glycine, citrate, lactate, nucleotides, and lipids, making it difficult to generalize findings across the different tumors [7]. There are also technical questions encountered while performing metabolomic analyses that may interfere in the characterization of a cancer metabolome, including variation between samples and sensitivity, particularly for extraction-dependent MS-based techniques [11].

### 3.2.3 Other applications of clinical metabolomics

Other applications of clinical metabolomics include nutrigenomics. One environmental factor important to consider refers to the consequences of food ingestion on human health. The role of many food components in the protection against nutritional disorders is also well documented [52–54]. The area of nutrigenomics has recently emerged with a particular focus on generating a picture of how gene expression changes when a human being is exposed to various nutrients [55–58]. This approach could allow the discovery of bioprotective foods. Strategies based on metabolomics are expected to play a pivotal role here based on the fact that nutrition by definition is aimed at maintaining cellular and organic homeostasis [59]. Moreover, it is known that some metabolic diseases or different types of cancer can be prevented by nutrition [60–67] and that many foods and drinks contain a rich variety of metabolites; thus measuring metabolites directly would be sensible. In this context, nutrigenomics can be expanded and should include tissues, cellular, or biofluid-specific nutritional metabolomes [59,68–71].

Finally, experimental pharmacology and toxicology differ from human nutrition in many aspects with regard to metabolomics. First, most of the investigations in toxicology and pharmacology are carried out in laboratory animals, which are genetically and nutritionally more homogeneous than humans. In addition, experiments in both toxicology and pharmacology involve the administration of a drug or a substance that will have a direct effect on the metabolism. Many drugs generally interact with a protein or enzyme receptor and finally, metabolic signals that act in a coordinated manner with the factors that regulate the disease have a marked effect on the determination of the metabolic profile. This directly affects the application of the metabolomics for the detection of different diseases, such as cardiovascular or neurological pathologies. Because of these differences, the signal-to-noise ratio will be higher in pharmacology and toxicology research than in human nutrition research. Thus it is clear that when it comes to human nutrition research, it is important to take into account the accuracy and

precision of metabolite measurements to ensure that the data obtained maintains the biological information that correlates with the phenotype of interest [72].

### 3.3 Techniques used in metabolomics and databases

Not all strategies used in metabolomics are universally accepted; however, it is possible to summarize the most popular ones [59,73–75] (see Table 3.1). Those techniques include target metabolite assay, targeted analysis, metabolite profiling, and metabolic fingerprinting. Table 3.1 includes a brief description of the goals that these strategies aim to achieve and which analytical approaches are commonly used.

The techniques that are useful for the evaluation of metabolomes include direct spectroscopy, chromatography (fingerprinting), or high-performance chromatography (GC or HPLC) combined with spectroscopy (IR, MS). The coupling of chromatographic methods with MS can substantially increase the depth of metabolome coverage and could also identify a more significant number of metabolites [76].

GC–MS is one of the most used analytical techniques in metabolomics. GC–MS is very useful for the analysis of a qualitatively and quantitatively wide range of volatile and/or derivatized nonvolatile metabolites with high thermal stability. Those metabolites are previously separated and the eluted compounds are identified by mass spectrophotometers. This technique has high analytical reproducibility and lower cost compared to other hyphenated techniques, such as liquid chromatography coupled to mass spectrometry (LC–MS) or LC–NMR [76,77]. Alternatively, direct injection MS analysis may also be applied for phenotyping of cell types or a physiological state of an organism, that is, FT-MS, which provides an ultimate limit of detection and mass measurement precision to enable metabolomic analyses [76,78].

Over the past few years, the application of LC–MS has also been growing up. LC–MS is more suitable than GC/MS for labile compounds and for those that are either hard to derivatize or render volatile. LC–MS has also been used in the nontargeted analysis of endogenous metabolites, which strives to detect, in an unbiased manner, as many metabolites as possible in the studied biological system [76,79–81]. A closely related method to LC–MS is capillary electrophoresis (CE)–MS [82]. CE coupled with MS is also a new type of analysis used in metabolite profiling with high resolution and reliability [83].

FT-IR, NIR, Raman, and more recently magnetic resonance (NMR) spectroscopy are continually developing rapid, nondestructive, reagent-less, and high-throughput techniques for a diverse range of sample types. Very recently, FT-IR has also been introduced as a metabolic fingerprinting technique to study defects of spermatogenesis and also within the plant sciences [84–86]. Raman spectroscopy coupled with microscopy has been recently used with great success, making possible the detection of many types of cancer; however, most of the applications are still far from clinical adoption [87].

NMR techniques require pure samples and a minimum quantity compared to other methods which need higher quantities of samples. NMR-based methods can be classified into solution NMR and insoluble or solid-state NMR, according to sample solubility. Using high-resolution-MS techniques, it is possible to acquire metabolic

**TABLE 3.1** Metabolomics strategies and common analytical platforms.

Strategy	Characteristics	Most used techniques
Targeted metabolite assay	Single assay of metabolites	Spectrophotometric HPLC/GC/LC–MS/GC–MS/NMR
Targeted analysis	Assay of multiple compounds of the same class	HPLC–MS/GC–MS/GC–TOF–MS/NMR
Metabolite profiling	Assay of specific compounds (i.e., amino acids, lipids, and organic acids) Sample extraction depends on the properties of each group of compounds	HPLC–MS/HILIC/CE–MS/HPLC–NMR/ LC–EC
Metabolite fingerprinting	Assay of complex profiles. Metabolites are not quantified individually Samples are analyzed directly or with very little extraction	DIMS/NMR/ESI–MSI/FT-IR

CE, Capillary electrophoresis; DIMS, direct infusion mass spectrometry; EC, electrochemical array; ESI, electrospray ionization; FT-IR, Fourier transform infrared spectroscopy; GC, gas chromatography; HILIC, hydrophobic interaction liquid chromatography; HPLC, high-performance liquid chromatography; MS, mass spectrometry; MSI, mass spectrometry imaging; NMR, nuclear magnetic resonance; TOF, time of fly.



fingerprints from insoluble and solid-state samples [88]. In one-dimensional NMR, protons ( $^1\text{H}$ ) are usually observed ( $^1\text{H}$ -NMR) due to the sensitivity and common occurrence of this magnetic nucleus. More detailed analyses, such as metabolite identification or flux assay, can be obtained with other nuclei, particularly  $^{13}\text{C}$  and  $^{15}\text{N}$  that are coupled with  $^1\text{H}$  nuclei in two-dimensional or multidimensional NMR analysis [89,90].

### 3.4 Conclusions

Metabolomics evidently has a high impact in areas, such as biology, that extend far beyond the scope of this chapter. The generation of high-throughput global data that will be used as part of the integrated methodologies will prove to be an expensive risk and will require extensive knowledge about computer modeling, physiology, and metabolism as well as excellent technical skills in measuring gene and protein expression and metabolic flux analysis. While the current trend of high-throughput data generation is increasingly popular, we believe that beneficial information could still be extracted from the data that has already been generated, as illustrated recently [91]. Such a multidisciplinary approach paves the way to establishing strong symbiotic research collaborations. On the whole, we can expect significant advances in the field of metabolic engineering as a consequence of elucidating previously unknown regulatory properties using such holistic approaches [92].

With the development of more sensitive and advanced instrumentation and computational tools for data interpretation in the physiological context, metabolomics has the potential to improve our understanding of the molecular mechanisms of disease. A state-of-the-art metabolomics analysis requires specialized training in areas related to biology, chemistry, physics, and bioinformatics. The adequate extraction of the samples is necessary to ensure an excellent quality of them. On the other hand, the improvement in automated metabolite identification, the establishment of spectral databases of metabolites, and novel computational tools to handle metabolomics data are also necessary [93].

The ultimate goal is to understand and to predict the behavior of complex systems (such as a whole organism) by using the results obtained from data mining tools for subsequent modeling and simulation. Taking into account the significant advances in other omic techniques, such as transcriptomics, genomics, and proteomics, the integrated analysis of these techniques with metabolomics and bioinformatics could be an excellent tool to identify and predict the function of genes and the subsequent alteration of metabolite levels, particularly those involved in human diseases [14].

### Acknowledgments

This work was supported by grants from ANPCyT (PICT 2014-2184, 2016-2188) and CONICET (PIP 476). MVB and DGC are research members from CONICET.

### References

- [1] Fiehn O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol Biol* 2002;48:155–71.
- [2] Putri SP, Yamamoto S, Tsugawa H, Fukusaki E. Current metabolomics: technological advances. *J Biosci Bioeng* 2013;116:9–16.
- [3] Hall IP. Pharmacogenetics, pharmacogenomics and airway disease. *Respir Res* 2002;3:10.
- [4] Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, et al. Potential of metabolomics as a functional genomics tool. *Trends Plant Sci* 2004;9:418–25.
- [5] Weckwerth W. Metabolomics in systems biology. *Annu Rev Plant Biol* 2003;54:669–89.
- [6] Kohler I, Verhoeven A, Derks RJ, Giera M. Analytical pitfalls and challenges in clinical metabolomics. *Bioanalysis* 2016;8:1509–32.
- [7] Griffin JL, Shockcor JP. Metabolic profiles of cancer cells. *Nat Rev Cancer* 2004;4:551–61.
- [8] Mendes P, Kell DB, Westerhoff HV. Channelling can decrease pool size. *Eur J Biochem* 1992;204:257–66.
- [9] Mendes P, Kell DB, Westerhoff HV. Why and when channelling can decrease pool size at constant net flux in a simple dynamic channel. *Biochim Biophys Acta* 1996;1289:175–86.
- [10] Boros LG, Lerner MR, Morgan DL, Taylor SL, Smith BJ, Postier RG, et al. [1,2- $^{13}\text{C}_2$ ]-D-Glucose profiles of the serum, liver, pancreas, and DMBA-induced pancreatic tumors of rats. *Pancreas* 2005;31:337–43.
- [11] Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* 2009;15:431–40.
- [12] Gomez-Casati DF, Zanol MI, Busi MV. Metabolomics in plants and humans: applications in the prevention and diagnosis of diseases. *BioMed Res Int* 2013;2013:792527.
- [13] Allwood JW, Ellis DI, Goodacre R. Metabolomic technologies and their application to the study of plants and plant-host interactions. *Physiol Plant* 2008;132:117–35.
- [14] Akiyama K, Chikayama E, Yuasa H, Shimada Y, Tohge T, Shinozaki K, et al. PRiMe: a web site that assembles tools for metabolomics and transcriptomics. *In Silico Biol* 2008;8:339–45.

- [15] Tebani A, Abily-Donval L, Afonso C, Marret S, Bekri S. Clinical metabolomics: the new metabolic window for inborn errors of metabolism investigations in the post-genomic era. *Int J Mol Sci* 2016;17.
- [16] Gomez-Casati DF, Grisolia MJ, Busi MV. The significance of metabolomics in human health. In: Kumar Da, Stylianou A, editors. *Medical and health genomics*. Academic Press; 2016. p. 89–100.
- [17] Imperlini E, Santorelli L, Orru S, Scolamiero E, Ruoppolo M, Caterino M. Mass spectrometry-based metabolomic and proteomic strategies in organic acidemias. *BioMed Res Int* 2016;2016:9210408.
- [18] Vernon HJ. Inborn errors of metabolism: advances in diagnosis and therapy. *JAMA Pediatr* 2015;169:778–82.
- [19] Chandler RJ, Zerfas PM, Shanske S, Sloan J, Hoffmann V, DiMauro S, et al. Mitochondrial dysfunction in mutant methylmalonic acidemia. *FASEB J* 2009;23:1252–61.
- [20] Richard E, Monteoliva L, Juarez S, Perez B, Desviat LR, Ugarte M, et al. Quantitative analysis of mitochondrial protein expression in methylmalonic acidemia by two-dimensional difference gel electrophoresis. *J Proteome Res* 2006;5:1602–10.
- [21] Gompertz D, Storrs CN, Bau DC, Peters TJ, Hughes EA. Localisation of enzymic defect in propionicacidemia. *Lancet* 1970;1:1140–3.
- [22] Wikoff WR, Nagle MA, Kouznetsova VL, Tsigelny IF, Nigam SK. Untargeted metabolomics identifies enterobiome metabolites and putative uremic toxins as substrates of organic anion transporter 1 (Oat1). *J Proteome Res* 2011;10:2842–51.
- [23] Mackenzie DY, Woolf LI. Maple syrup urine disease; an inborn error of the metabolism of valine, leucine, and isoleucine associated with gross mental deficiency. *Br Med J* 1959;1:90–1.
- [24] Frazier DM, Allgeier C, Homer C, Marriage BJ, Ogata B, Rohr F, et al. Nutrition management guideline for maple syrup urine disease: an evidence- and consensus-based approach. *Mol Genet Metab* 2014;112:210–17.
- [25] Miller MJ, Kennedy AD, Eckhart AD, Burrage LC, Wulff JE, Miller LA, et al. Untargeted metabolomic analysis for the clinical screening of inborn errors of metabolism. *J Inher Metab Dis* 2015;38:1029–39.
- [26] Pie J, Lopez-Vinas E, Puisac B, Menao S, Pie A, Casale C, et al. Molecular genetics of HMG-CoA lyase deficiency. *Mol Genet Metab* 2007;92:198–209.
- [27] Kolker S, Christensen E, Leonard JV, Greenberg CR, Boneh A, Burlina AB, et al. Diagnosis and management of glutaric aciduria type I—revised recommendations. *J Inher Metab Dis* 2011;34:677–94.
- [28] Popek M, Walter M, Fernando M, Lindner M, Schwab KO, Sass JO. Two inborn errors of metabolism in a newborn: glutaric aciduria type I combined with isobutyrylglycinuria. *Clin Chim Acta: Int J Clin Chem* 2010;411:2087–91.
- [29] McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;9:356–69.
- [30] Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443–53.
- [31] Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007;39:857–64.
- [32] Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008;40:638–45.
- [33] Gieger C, Geistlinger L, Altmaier E, Hrabce de Angelis M, Kronenberg F, Meitinger T, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. *PLoS Genet* 2008;4:e1000282.
- [34] Kathiresan S, Melander O, Guiducci C, Surti A, Burt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008;40:189–97.
- [35] Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008;40:161–9.
- [36] Fossel ET, Carr JM, McDonagh J. Detection of malignant tumors. Water-suppressed proton nuclear magnetic resonance spectroscopy of plasma. *N Engl J Med* 1986;315:1369–76.
- [37] Denkert C, Budczies J, Kind T, Weichert W, Tablack P, Sehouli J, et al. Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. *Cancer Res* 2006;66:10795–804.
- [38] Glunde K, Serkova NJ. Therapeutic targets and biomarkers identified in cancer choline phospholipid metabolism. *Pharmacogenomics* 2006;7:1109–23.
- [39] Kline EE, Treat EG, Averna TA, Davis MS, Smith AY, Sillerud LO. Citrate concentrations in human seminal fluid and expressed prostatic fluid determined via <sup>1</sup>H nuclear magnetic resonance spectroscopy outperform prostate specific antigen in prostate cancer detection. *J Urol* 2006;176:2274–9.
- [40] Morvan D, Demidem A. Metabolomics by proton nuclear magnetic resonance spectroscopy of the response to chloroethylnitrosourea reveals drug efficacy and tumor adaptive metabolic pathways. *Cancer Res* 2007;67:2150–9.
- [41] Serkova NJ, Spratlin JL, Eckhardt SG. NMR-based metabolomics: translational application and treatment of cancer. *Curr Opin Mol Ther* 2007;9:572–85.
- [42] Bathen TF, Jensen LR, Sitter B, Fjosne HE, Halgunset J, Axelson DE, et al. MR-determined metabolic phenotype of breast cancer in prediction of lymphatic spread, grade, and hormone status. *Breast Cancer Res Treat* 2007;104:181–9.
- [43] Ackerstaff E, Glunde K, Bhujwalla ZM. Choline phospholipid metabolism: a target in cancer cells? *J Cell Biochem* 2003;90:525–33.
- [44] Mazurek S, Eigenbrodt E. The tumor metabolome. *Anticancer Res* 2003;23:1149–54.
- [45] Bathen TF, Engan T, Krane J, Axelson D. Analysis and classification of proton NMR spectra of lipoprotein fractions from healthy volunteers and patients with cancer or CHD. *Anticancer Res* 2000;20:2393–408.
- [46] Bolan PJ, Meisamy S, Baker EH, Lin J, Emory T, Nelson M, et al. In vivo quantification of choline compounds in the breast with <sup>1</sup>H MR spectroscopy. *Magn Reson Med* 2003;50:1134–43.
- [47] Jacobs MA, Barker PB, Bottomley PA, Bhujwalla Z, Bluemke DA. Proton magnetic resonance spectroscopic imaging of human breast cancer: a preliminary study. *J Magn Reson Imaging* 2004;19:68–75.
- [48] Stanwell P, Gluch L, Clark D, Tomanek B, Baker L, Giuffre B, et al. Specificity of choline metabolites for in vivo diagnosis of breast cancer using <sup>1</sup>H MRS at 1.5 T. *Eur Radiol* 2005;15:1037–43.

- [49] Yeung DK, Cheung HS, Tse GM. Human breast lesions: characterization with contrast-enhanced in vivo proton MR spectroscopy—initial results. *Radiology* 2001;220:40–6.
- [50] Erben V, Bhardwaj M, Schrotz-King P, Brenner H. Metabolomics biomarkers for detection of colorectal neoplasms: a systematic review. *Cancers* 2018;10:E246.
- [51] Yu L, Li K, Zhang X. Next-generation metabolomics in lung cancer diagnosis, treatment and precision medicine: mini review. *Oncotarget* 2017;8:115774–86.
- [52] Davis CD, Milner J. Frontiers in nutrigenomics, proteomics, metabolomics and cancer prevention. *Mutat Res* 2004;551:51–64.
- [53] Afman L, Muller M. Nutrigenomics: from molecular nutrition to prevention of disease. *J Am Diet Assoc* 2006;106:569–76.
- [54] Fergusson LR. Nutrigenomics—integrating genomic approaches into nutrition research. *Mol Diagn Ther* 2006;10:101–8.
- [55] Mariman ECM. Nutrigenomics and nutrigenetics: the \_omics\_ revolution in nutritional science. *Biotechnol Appl Biochem* 2006;44:119–28.
- [56] Mutch DM, Wahli W, Williamson G. Nutrigenomics and nutrigenetics: the emerging faces of nutrition. *FASEB J* 2005;19:1602–16.
- [57] Ordoas JM, Mooser V. Nutrigenomics and nutrigenetics. *Curr Opin Lipidol* 2004;15:101–8.
- [58] Roche HM. Nutrigenomics—new approaches for human nutrition research. *J Sci Food Agric* 2006;86:1156–63.
- [59] Goodacre R. Metabolomics of a superorganism. *J Nutr* 2007;137:259S–66S.
- [60] Berna G, Oliveras-Lopez MJ, Jurado-Ruiz E, Tejedo J, Bedoya F, Soria B, et al. Nutrigenetics and nutrigenomics insights into diabetes etiopathogenesis. *Nutrients* 2014;6:5338–69.
- [61] Braicu C, Mehterov N, Vladimirov B, Sarafian V, Nabavi SM, Atanasov AG, et al. Nutrigenomics in cancer: revisiting the effects of natural compounds. *Semin Cancer Biol* 2017;46:84–106.
- [62] Corella D, Ordoas JM. Nutrigenomics in cardiovascular medicine. *Circ Cardiovasc Genet* 2009;2:637–51.
- [63] Fergusson LR. Dissecting the nutrigenomics, diabetes, and gastrointestinal disease interface: from risk assessment to health intervention. *Omics* 2008;12:237–44.
- [64] Kang JX. Nutrigenomics and cancer therapy. *J Nutrigenet Nutrigenomics* 2013;6:1–I.
- [65] Ardekani AM, Jabbari S. Nutrigenomics and cancer. *Avicenna J Med Biotechnol* 2009;1:9–17.
- [66] Trujillo E, Davis C, Milner J. Nutrigenomics, proteomics, metabolomics, and the practice of dietetics. *J Am Diet Assoc* 2006;106:403–13.
- [67] Virgili F, Perozzi G. How does nutrigenomics impact human health? *IUBMB Life* 2008;60:341–4.
- [68] Kaput J, Rodriguez RL. Nutritional genomics: the next frontier in the postgenomic era. *Physiol Genom* 2004;16:166–77.
- [69] van Ommen B. Nutrigenomics: exploiting systems biology in the nutrition and health arenas. *Nutrition* 2004;20:4–8.
- [70] van Ommen B, Stierum R. Nutrigenomics: exploiting systems biology in the nutrition and health arenas. *Curr Opin Biotechnol* 2002;13:517–21.
- [71] Zeisel SH, Freake HC, Bauman DE, Bier DM, Burrin DG, German JB, et al. The nutritional phenotype in the age of metabolomics. *J Nutr* 2005;135:1613–16.
- [72] Gibney MJ, Walsh M, Brennan L, Roche HM, German B, van Ommen B. Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr* 2005;82:497–503.
- [73] Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics* 2001;2:155–68.
- [74] Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* 2004;22:245–52.
- [75] Harrigan GG, Goodacre R. *Metabolic profiling: its role in biomarker discovery and gene function analysis*. Kluwer Academic Publishers; 2003.
- [76] Bedair M, Sumner L. Current and emerging mass-spectrometry technologies for metabolomics. *Trends Anal Chem* 2008;27:238–50.
- [77] Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L. Metabolite profiling for plant functional genomics. *Nat Biotechnol* 2000;18:1157–61.
- [78] Aharoni A, Ric de Vos CH, Verhoeven HA, Maliepaard CA, Kruppa G, Bino R, et al. Nontargeted metabolome analysis by use of Fourier transform ion cyclotron mass spectrometry. *Omics* 2002;6:217–34.
- [79] De Vos RC, Moco S, Lommen A, Keurentjes JJ, Bino RJ, Hall RD. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat Protoc* 2007;2:778–91.
- [80] Moco S, Bino RJ, Vorst O, Verhoeven HA, de Groot J, van Beek TA, et al. A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiol* 2006;141:1205–18.
- [81] Moco S, Capanoglu E, Tikunov Y, Bino RJ, Boyacioglu D, Hall RD, et al. Tissue specialization at the metabolite level is perceived during the development of tomato fruit. *J Exp Bot* 2007;58:4131–46.
- [82] Dunn WB, Bailey NJ, Johnson HE. Measuring the metabolome: current analytical technologies. *Analyst* 2005;130:606–25.
- [83] Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J Proteome Res* 2003;2:488–94.
- [84] Gilany K, Pouracil RS, Sadeghi MR. Fourier transform infrared spectroscopy: a potential technique for noninvasive detection of spermatogenesis. *Avicenna J Med Biotechnol* 2014;6:47–52.
- [85] Kaderbhai NN, Broadhurst DI, Ellis DI, Goodacre R, Kell DB. Functional genomics via metabolic footprinting: monitoring metabolite secretion by *Escherichia coli* tryptophan metabolism mutants using FT-IR and direct injection electrospray mass spectrometry. *Comp Funct Genomics* 2003;4:376–91.
- [86] Baranska M, Schulz H, Rosch P, Strehle MA, Popp J. Identification of secondary metabolites in medicinal and spice plants by NIR-FT-Raman microspectroscopic mapping. *Analyst* 2004;129:926–30.
- [87] Santos IP, Barroso EM, Bakker Schut TC, Caspers PJ, van Lanschot CGF, Choi DH, et al. Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics. *Analyst* 2017;142:3025–47.
- [88] Bertocchi F, Paci M. Applications of high-resolution solid-state NMR spectroscopy in food science. *J Agric Food Chem* 2008;56:9317–27.



- [89] Kikuchi J, Shinozaki K, Hirayama T. Stable isotope labeling of *Arabidopsis thaliana* for an NMR-based metabolomics approach. *Plant Cell Physiol* 2004;45:1099–104.
- [90] Sekiyama Y, Kikuchi J. Towards dynamic metabolic network measurements by multi-dimensional NMR-based fluxomics. *Phytochemistry* 2007;68:2320–9.
- [91] De Backer MD, Thurmond RL, Carmen AA, Luyten WH. Gene-expression-based responses to drug treatment. *Drug News Perspect* 2002;15:155–65.
- [92] Vemuri GN, Aristidou AA. Metabolic engineering in the -omics era: elucidating and modulating regulatory networks. *Microbiol Mol Biol Rev* 2005;69:197–216.
- [93] Kouskoumvekaki I, Panagiotou G. Navigating the human metabolome for biomarker identification and design of pharmaceutical molecules. *J Biomed Biotechnol* 2011;2011:525497.



# Clinical applications of next-generation sequencing

*IAN Tully*

Institute of Medical Genetics, Division of Cancer & Genetics, University Hospital of Wales, Cardiff, United Kingdom

## 4.1 Introduction

The practice of medical genetics has been transformed in recent years by the increasing availability of next-generation sequencing (NGS) as a diagnostic and research tool. Classically genetic investigations can be thought of in two streams; cytogenetics and molecular genetics, with cytogenetics focusing on chromosomal structure and molecular genetics on gene sequencing. In the last 5 years, molecular genetic testing has changed completely, from an approach where single genes would be chosen for sequencing based on a patient's phenotype, to an approach where a large number of genes, or even the whole genome can be sequenced simultaneously.

NGS refers to any technology in which large amounts of genetic material are simultaneously sequenced—this is also referred to as high-throughput sequencing or massively parallel sequencing. NGS can include the sequencing of whole-genome sequencing (WGS) and whole-exome sequencing (WES), where the coding regions of genes are sequenced, or large gene panels where specific genes of relevance to the phenotype being investigated are captured and sequenced.

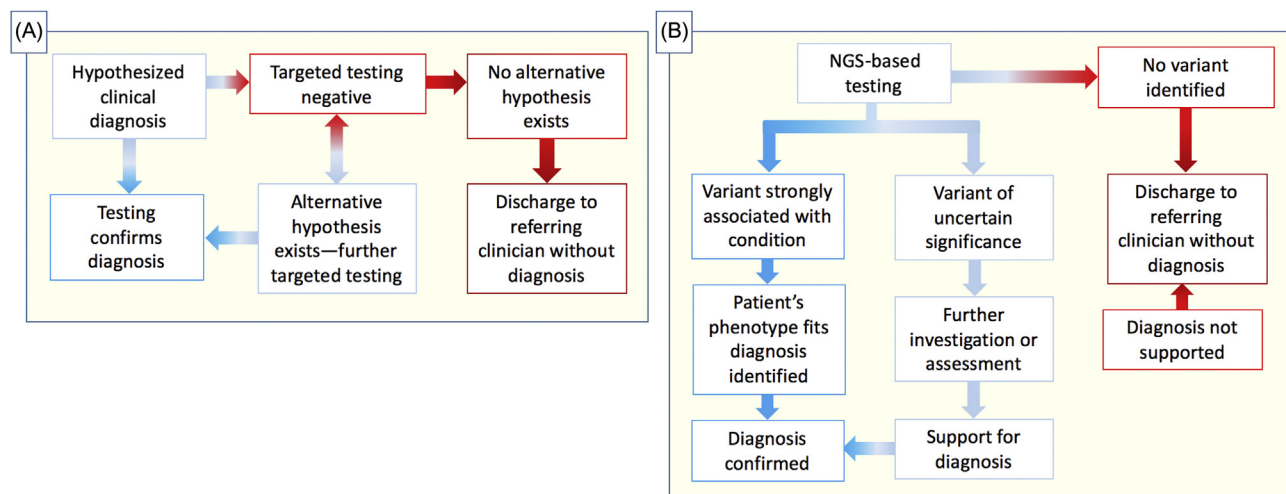
NGS has been referred to as a “hypothesis-free” investigation. In reality the practice of medical genetics now involves the synthesis of genomic data alongside clinical data to reach a diagnostic hypothesis, which can then be confirmed by reassessment of the patient or by further, more routine investigations.

For example, a child with developmental delay and subtle but undistinctive dysmorphic facial features and a normal chromosomal microarray (CMA) may be investigated by NGS. A candidate variant may be found for a condition in which abnormal phalangeal appearance on X-ray is pathognomonic for the condition. A radiograph could then be arranged to provide confirmatory diagnostic information that may otherwise have been missed.

In another case a patient could have a striking set of clinical features, but the pattern goes unrecognized by the geneticist. WGS could identify a variant associated with a rare condition that the practitioner was not previously aware of. On reassessment of the patient, their features are found to be a close match for this rare condition, and the diagnosis can be confirmed.

The field of genetic research has also been transformed by the advent of high-throughput sequencing. In an example relevant to pediatrics, the deciphering developmental disorders project involves over 12,000 patients with previously undiagnosed developmental disorders and used sequential testing with CMA, WES, and WGS to look for diagnoses. The project has not only led to limited diagnoses being made thus far, but the discovery of new disease genes and the expansion of the described phenotype for many conditions have made huge impact in clinical practice.

The advent of a transformative technology to any field of study presents challenges for the adherents to that field. Even for full-time medical geneticists and genetic counselors, keeping up with the rate of progress in genomic medicine is challenging—let alone for health professionals working in other specialties. Despite the challenges presented, the potential benefits for patient care in terms of improved diagnostics and, eventually, personalized medicine cannot be overstated. Indeed, the use of NGS analysis of tumor samples has already



**FIGURE 4.1** Schematic of investigations undertaken in medical genetics (A) pre-NGS era and (B) with advent of readily available NGS testing. NGS, Next-generation sequencing.

allowed for oncologists to personalize chemotherapeutic regimes to a degree not previously possible, with improved outcomes both in terms of side effects and remission rates.

In this chapter, we will explore the NGS technologies currently available in clinical practice, along with their advantages and disadvantages. We will consider the appropriate use of these technologies and the ethical implications of NGS testing. We will review the process of generating clinically useful information from raw sequence data and the clinical interpretation of variants discovered using NGS. Finally, we will examine emerging trends and possible future developments in sequencing technology and their potential benefits to clinical practice and biomedical research (Fig. 4.1).

## 4.2 Next-generation sequencing technologies

To understand the data generated by NGS platforms, and to appreciate some of its limitations, it is helpful to understand the processes by which the data is acquired. To understand NGS, it is helpful to understand Sanger sequencing.

### 4.2.1 First generation—Sanger sequencing

Fred Sanger's chain termination method involved the incorporation of 3' dead dideoxynucleotides into a DNA sequence. These nucleotides lack the hydroxyl group necessary for incorporation of the next nucleotide in a sequence. By including a low concentration of a pool of normal nucleotides used for a DNA polymerase reaction based on a template strand, sequences of every possible length are produced. Using parallel reactions with radiolabeled dideoxynucleotides of one type in each reaction, it is possible to identify the last base incorporated into each sequence. The sequences can then be separated using chromatography or electrophoresis before imaging takes place.

Various refinements to Sanger's technique were made over the years, offering greater fidelity and ease of interpretation, including fluorescently labeled nucleotides and capillary electrophoresis, generating a sequence by the detection of different colored fluorophores as DNA molecules passed down a glass tube.

The underlying principal of sequencing by synthesis is fundamental to many NGS techniques.

### 4.2.2 Second- (next-) generation sequencing

When a nucleotide is incorporated into a DNA molecule, pyrophosphate is released. Pyrosequencing utilizes an enzyme to convert pyrophosphate into ATP, which in turn is used by Luciferase to catalyze the breakdown of luciferin, creating a small pulse of light, which can be read by a detector.

Single-stranded DNA molecules, bound to a stable surface, are washed in a solution containing a single nucleotide, DNA polymerase, sulfurylase, luciferin, and luciferase. The process is repeated thousands of times with different nucleotides. Each time a nucleotide is incorporated, a light pulse is detected and the base incorporated is recorded.

By combining the pyrosequencing approach with improvements in miniaturization and computing, 454 life technologies designed the first NGS platform. DNA molecules are captured and amplified on beads, which are then washed across a sequencing cell containing (wells). The wash procedure described earlier is then used to perform the sequencing off all strands simultaneously. It is this massive parallelization that allows for the rapid generation of whole-genome scale data.

Next significantly influential development was Illumina's solexa sequencing. The DNA to be sequenced is prepared by ligation of adapters. These adapters are complementary to adapters on the flow cell, which bind the DNA. The DNA molecules then arch over to bind with the adapter on the other end, before sequencing takes place. This process is repeated multiple times to generate localized areas with high concentrations of the same DNA sequence.

The method uses the incorporation of fluorescent reversible-terminator nucleotides, which do not allow more than one nucleotide to be incorporated at once but can be enzymatically converted to normal nucleotides after the signal is read. Similar to pyrosequencing, the nucleotides are washed over the strands with DNA polymerase in cycles. Between each cycle, the strands are excited with lasers and the fluorophores read. An enzyme solution, to remove the blocking molecules, is then passed over the strands before the next nucleotide wash.

### 4.2.3 The third-generation sequencing

For the purpose of this text, third-generation sequencing refers to techniques in which the DNA molecule is read directly, without the need for amplification. This single-molecule sequencing approach results in very long reads—up to 10 kB have been reported with both the Oxford Nanopore Technology (ONT) minION and with the Pacific Biosciences (PACBIO) Single Molecule Real-Time (SMRT) cell platform.

Third-generation sequencing is still in its infancy and is used as a research tool, rather than in clinical practice; however, with the speed at which developments in the field of genomics are occurring, it is certainly feasible to imagine that clinical applications will be found.

ONT uses a protein pore, embedded in a flow cell membrane, through which single-stranded DNA passes. The DNA is prepared by binding to library proteins that anchor it on the membrane of the flow cell next to the pores, so that the molecules can be captured and fed in. Each pentamer that passes through the pore generates a different micro-voltage, which is read to determine the DNA sequence. Each nucleotide is read five times as part of sequential pentamers.

At present, nanopore technology is only able to offer an accuracy rate of around 90%, increasing to 94% with more recent flow cells in model organisms. Part of the inaccuracy arises from the tendency of DNA strands not to pass through the pore at a constant speed, leading to false calling of insertions and deletions. The other main source of inaccuracy is the difficulty in distinguishing between minute fluctuations in the micro-current between pentamers that are close in composition.

The PACBIO technology uses zero-mode waveguides, pores narrower than the wavelength of light. When light is emitted through these pores, it does not reach the end, creating what the company describes as the world's smallest light microscope. The DNA molecule to be sequenced is held at the base of the bore by an enzyme, while bathed in fluorophore-labeled nucleotides. As these nucleotides are incorporated, the fluorophore is cleaved and detected.

Although the PACBIO offers greater accuracy than the Oxford Nano Pore (ONP) platform, the relative cost and footprint make it difficult to adopt on a clinical basis at present.

---

## 4.3 Choice of test

Despite the promised wonders of WGS, it may not be the most best choice of investigation for every patient. There are several factors to consider when choosing the most appropriate genomic investigation for your patient.

Most important from a technical point of view, the question is whether sequencing will actually give you the information you need. Not every genetic diagnosis is caused by a change in the coding sequence of the genome, and not every sequence change will be detected by NGS.

Triplet repeat expansions have been the *bête-noire* of molecular geneticists since sequencing was first implemented. These disorders are caused when normally existing tracts of repeating codons expand beyond their normal limits. In some cases, this expansion disrupts gene transcription, such as in fragile X syndrome, Friedreich's ataxia, and myotonic dystrophy. In other cases a toxic protein product is made, such as in the accumulation of neurotoxic glutamate tracts in Huntington's disease and several of the inherited spinocerebellar ataxia syndromes.

As discussed earlier, the second-generation sequencing technologies that form the backbone of clinical genomics struggle to accurately determine the length of repetitive tracts of DNA, leading to both false negative and false positive calls.

Some genetic conditions are caused not by a mutation in the protein-coding sequence, but by changes in the regulation of gene transcription. DNA methylation at cytosine guanine (CG) residue-rich regions is widely understood to downregulate DNA transcription at these sites. Multiple diseases are known to be associated with disruption to methylation. If a patient has features that are in keeping with a diagnosis known to be caused by abnormal methylation, for example, Prader–Willi syndrome or Angelman syndrome then NGS is likely to miss the responsible defect and, with it, the diagnosis.

The cost of the chosen investigation, along with the pretest probability of finding a diagnosis, should also be considered when choosing an investigation. Although the cost of WES is now well below £1000 a patient in multiplexed runs, trio sequencing is often necessary to identify *de novo* candidate variants for disease. When sequencing three (or more) members of the family is necessary for full interpretation, the costs can mount.

In general, for patients with undiagnosed developmental delay and negative CMA, WES successfully identifies a genetic cause for the condition in around 30% of the cases. In similar cohorts, WGS is effective in around 40% of cases. Large gene panels, curated to include as many Online Mendelian Inheritance in Man (OMIM) pathological genes, have a similar efficacy to WES, and a lower price point.

These numbers are estimates based on current published data and may improve as clinicians become more adept at choosing the best candidates for NGS testing and at the interpretation of variants.

We will consider the appropriate use of small gene panels, large gene panels, WES and WGS, along with the advantages and disadvantages of each approach.

#### 4.3.1 *Small gene panels*

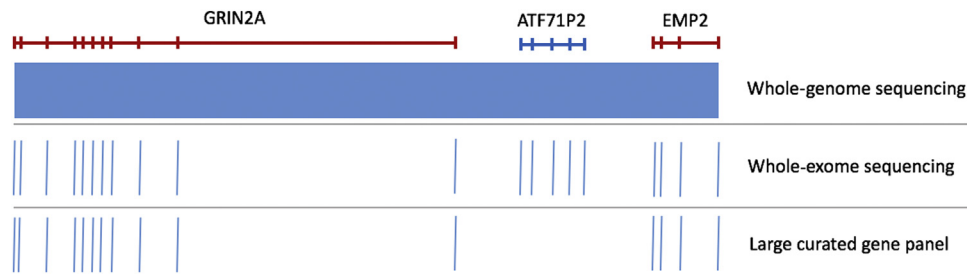
In cases where the phenotype is obvious and uncomplicated, the depth of analysis offered by WES can be considered excessive—the equivalent of ordering a whole-body MRI scan to investigate a headache with symptoms of increased intracranial pressure instead of a targeted MRI brain.

The amount of information to filter increases the risk of the diagnosis being missed. The deluge of information also raises the risk of variants of uncertain significance (VUS) in important genes being identified, obligating further investigations without the promise of answers and causing unnecessary anxiety for the patient and their family.

Fortunately, a large number of well-curated gene panels are available. The UK Genetic Testing Network (UKGTN) keeps an up-to-date list of panels offered by different centers around the United Kingdom. The design and interpretation of these panels is highly specialist, and the panels offered at different centers reflect the local expertise and experience at those units. The information found on UKGTN includes the cost, turnaround time, and a list of the genes included in the analysis, as well as links to the OMIM website where the phenotypes and evidence linking the gene to the phenotype are described in detail.

Before ordering a gene panel, it is crucial to ensure that any genes you want sequencing are actually included in the panel. To use a fairly common example, there are multiple retinal degeneration panels available online—the first two hits on UKGTN are a 55-gene panel at a cost of £650 and a 105-gene panel for £897. Prior to making the decision about which to go for, a clinician will have to review the differences between the panels and decide if the extra expense is valid.

To consider the question in more detail, a review of the panels shows that the larger panel has more forms of Bardet–Biedl syndrome (BBS) included. BBS has a striking phenotype beyond the retinal problems, including polydactyly, obesity, and developmental delay. If the patient does not have additional features of BBS then these extra genes being sequenced is not likely to provide any greater diagnostic insight.



**FIGURE 4.2** Schematic overview of a region of chromosome 16p. GRIN2A gene associated with Landau–Kleffner syndrome (aphasic epilepsy and speech delay), ATF71P2 gene not currently associated with any diagnosis, EMP2 gene associated with nephrotic syndrome type 10. Whole-genome sequencing comprises all material, including introns and intragenic DNA. WES includes all coding exons. LCP includes only exons of genes known to be associated with disease at the time of panel design. *LCPs*, Large curated gene panels; *WES*, whole-exome sequencing.

Examples of the type of disorders for which gene panels may be appropriate include retinopathy, sensorineural hearing loss, aortic dissection, epidermolysis bullosa, or any simplex phenotype.

Panels are also of use for more complex phenotypes in centers, which have not yet made the transition to the use of WES as a baseline test. In some centers, where funding for WES may be more limited, a gene panel may be the first line of investigation with WES only being considered where a diagnosis was *not* reached on panel (Fig. 4.2).

#### 4.3.2 Whole-exome sequencing and large curated panels

Illumina offers two large curated gene panels (LCPs): the TruSight One featuring 4813 genes (12 Mb genomic material) and the TruSight One Expanded featuring 6794 genes (16.6 Mb genomic material). They are curated to include disease-associated genes based on published research. The Expanded panel is a more recent addition and includes disease-related genes that were not known at the time the original panel was designed.

WES covers 45 Mb of genomic material, over 98% of the coding genome. The techniques and use of these LCPs overlaps considerably with WES, indeed LCPs are sometimes referred to as clinical exomes, and WES can also be thought of as a curated gene panel covering all known OMIM genes. We will therefore consider them together.

In both cases a vast amount of data is generated, and the pipeline for processing the data must be considered. There are two basic approaches; proband sequencing and trio sequencing. In trio sequencing the proband's parents undergo the same investigation—this helps to distinguish variants that have been inherited from *de novo* variants, which is essential when considering heterozygous missense mutations. In proband sequencing the proband is sequenced alone, with segregation analysis of identified variants performed after the results. Proband sequencing may be the only option in cases where both parents are not available due to estrangement or death.

Advantages of trio sequencing include improved variant filtering before clinical interpretation, which is explored more in the sections on bioinformatics and clinical interpretation. It has the obvious disadvantage of being three times the cost of proband sequencing.

In general, WES and LCP sequencing are of greatest use in phenotypes where there is very significant overlap between a large number of different genetic causes. Developmental delay is a classic example, and there has been much research into the use of WES in making a diagnosis in developmental delay. At most recent review the 100,000 genomes project (100kGP) lists 1927 genes as being linked to developmental delay, with more being added frequently.

The American College of Medical Genetics and Genomics (ACMG) suggests the following indications for WES/WGS: (Directors, 2012 no. 2658)

1. The phenotype or family history data strongly implicate a genetic etiology, but the phenotype does not correspond with a specific disorder for which a genetic test targeting a specific gene is available on a clinical basis.
2. A patient presents with a defined genetic disorder that demonstrates a high degree of genetic heterogeneity, making WES or WGS analysis of multiple genes simultaneously a more practical approach.
3. A patient presents with a likely genetic disorder, but specific genetic tests available for that phenotype have failed to arrive at a diagnosis.



4. A fetus with a likely genetic disorder in which specific genetic tests, including targeted sequencing tests, available for that phenotype have failed to arrive at a diagnosis.

Point (4) comes with a caveat that at present, turnaround time for WES is not sufficient for use in a prenatal setting, where urgent decisions are required about whether to proceed with the pregnancy.

Several recent publications have highlighted the utility of WES as a “first-tier” genetic test in the investigation of pediatric patients, alongside CMA and fragile X in boys; however, practice in this regard still varies according to local guidelines. In some centers, WES and LCP sequencing are considered only for patients who have undergone other tests without receiving a diagnosis. Although it is difficult to predict the future, it is certainly possible to imagine a scenario in a few years’ time where WES is the first genetic test offered to patients where the clinical diagnosis is not immediately clear.

---

## 4.4 Whole-genome sequencing

WGS covers all 3 Gb of the human genome, including the majority of intronic and exonic materials. There are technical and pathological reasons why WGS may be a better test than WES.

On a technical level the coverage offered by WGS is more even across the exome than in WES. WES relies upon an enrichment phase to capture the exonic material for sequencing; different exons are captured with differing rates of success. Similarly, different exons amplify with different efficiencies. This leads to artifactual fluctuations in the coverage between exons and genes, which is difficult to interpret, making calling copy number variations (CNVs) and structural variants from WES unreliable.

WGS suffers from no such variation. In WGS capture proceeds more or less evenly across all fragments of the genome as it is performed in a sequence naive fashion. Because the efficiency of the capture reaction is not dependant on sequence-specific primers, differences between the primer sequences cannot lead to differences in efficiency.

The even coverage means that WGS can be used to call both large and small CNVs and structural variants, allowing for the detection of mutations that could not previously be found. CMA has a resolution of around 100 kb for the detection of deletions and duplications. CNVs upto 100 kb can be identified by WGS, leading to diagnoses that could not otherwise be made.

It is currently believed that intragenic and intronic mutations can cause disease by interfering with binding sites for regulatory elements or disturbing the 3D architecture of the genome. These nonexonic variants are not well understood enough to be used in routine diagnosis at present; however, theoretically WGS is able to identify them.

Despite its advantages, WGS is currently prohibitively expensive for use in routine clinical practice. The practical aspect of processing the vast quantity of data generated as well as the issue of storing such large data files also presents challenges that need to be met before the technology can be more widely adopted.

---

## 4.5 Summary

Clinical use of NGS is still developing, and the information presented here represents a snapshot of a rapidly progressing field. At present, despite the evidence for its increased diagnostic yield and technical superiority, the cost, time intensity, manpower demands and informatic infrastructure obstacles prevent WGS being adopted for the majority of clinical uses. Its use as a research tool is better established.

In the medical genetics clinic, gene panels and WES are the mainstay of NGS investigations currently being offered. Their diagnostic yield is significantly greater than CMA and sequential single-gene analysis, but in time they are likely to be superseded by WGS once this becomes practical to implement.

---

## 4.6 Ethical considerations

Clinical geneticists have always had to confront the following key ethical issues when considering genetic testing:

1. VUS
2. Incidental findings (IFs)

3. When to offer genetic testing to patients unable to consent—this is, of course, of particular relevance to the practice of pediatrics.

VUS are a common finding in the practice of medical genetics. Every genome contains rare variants, with an average of >250 potentially disruptive exonic variants expected in each genome sequenced. For a number of reasons the impact of a variant on the translated protein, or the person's phenotype, is not always clear.

The interpretation of variants is covered in more detail in the section on clinical interpretation of genomic data. In this section, we will explore the process of dealing with uncertainty.

Some parents on being told of VUS are able to accept uncertainty, whereas others find such results to be a source of significant anxiety and stress. The impact of stress on a person's health and their family's well-being should not be underestimated, and therefore an uncertain result can be considered a potentially harmful outcome of genetic testing.

In some cases, families will falsely interpret VUS as negative results, in others they may falsely interpret VUS as positive. Most clinical geneticists have had the experience of VUS becoming codified as a diagnosis in a patient's subsequent medical notes. It goes without saying that making health-care decisions based on information that has not been understood can have negative consequences.

The chance of finding VUS has to be included in the process of consent. If a family feels that dealing with an uncertain result would be very difficult for them then testing may have to be delayed while they have more time to consider their options.

IFs are findings relevant to a patient's health, but not relevant to the presenting complaint. For example, a WES performed as a trio to identify the cause of developmental delay in a child may identify a pathogenic mutation in the Troponin T (TNNT2) gene in the father, which with cardiomyopathy in the father. Incidental findings are expected to occur in 2%–4% of WES performed.

In some cases, IFs will identify conditions that require further investigation or management. The father with the TNNT2 mutation would require an echocardiogram, ECG, and cardiology appointment to assess his left ventricular function. He may require medical management and significant lifestyle changes to ameliorate his risk of heart failure or sudden cardiac death.

While the early diagnosis of treatable conditions is widely acknowledged as beneficial, not all incidental findings will be related to treatable conditions. Imagine if, instead of a variant in TNNT2, the father was found to have a pathogenic variant in presenilin-1 (PSEN1). PSEN1 mutations are associated with early-onset Alzheimer's disease, presenting before the age of 65 years and sometimes even before 55 years of age.

At the time of writing, Alzheimer's disease (early-onset or otherwise) remains stubbornly resistant to treatment. Early detection in Alzheimer's is therefore of little direct benefit at this time, and receiving such a diagnosis could have a significant negative impact on a patient's emotional well-being for no long-term benefit.

The situation is complicated further when that father's age is considered. At 25 years old, he probably has around 30 years before the onset of the condition. In the next 30 years, there will be advances in medicine that we probably cannot even imagine at this point, and it is possible that Alzheimer's will be treatable, if diagnosed in a preclinical state. Withholding the knowledge therefore has a potential direct negative impact.

When offering presymptomatic genetic testing for incurable neurological diseases, it is a standard practice to give patients time to consider their decision after the first appointment. Many patients who attend the first appointment are keen on testing; but when asked to think about the real implications of early diagnosis, they do not attend the second one. An incidental finding of such a condition does not allow for the prediagnosis counseling felt necessary to prepare a patient to receive such information, so it is possible that the potential of negative psychological impact is even greater when the finding is entirely unexpected.

In recognition of the significant ethical challenges presented by IFs, the ACMG has published a "minimum list" of genes in which pathogenic variants should always be reported. This list includes multiple cancer predispositions, arrhythmogenic conditions, cardiomyopathies, hypercholesterolemia, and malignant hyperthermia. It does not include untreatable conditions.

In publication of the list the authors acknowledge that the perspective is likely to evolve as the clinical use of NGS evolves. The list, as with this chapter, exists in a transitional moment as the first attempt to come in terms with a unique set of challenges for which existing ethical norms are sometimes impractical.

The importance of the informed consent for genetic testing cannot be overstated, and this need is perhaps even more acute when considering NGS testing. The ACMG guidelines acknowledge that a tension exists between what is practical and what is most ethically correct.

For example, to obtain fully informed consent for WES and reporting of the ACMG minimal list, time should be taken to talk through each of the conditions on the list and answer any questions the patients may have about follow-up, and other implications should a pathogenic variant to be found in any one gene. Even in the hands of an experienced genetic counselor, this conversation could take several hours. This is not only impractical from a time management perspective, the hope of a patient engaging with such a detailed discussion is slight. As WES becomes more mainstream and is ordered by physicians with less genetic counseling experience, the problem only becomes more significant.

Instead a brief overview of the process and of the types of conditions in which IFs will be reported is recommended. It is acknowledged that this somewhat diminishes the informed consent process and deprives the patients of a degree of autonomy in making the decisions. Many consent protocols will ask the patient whether they wish to receive results pertaining to disorders on the minimal list and exclude these genes from the analysis if the patient answers no. Research into population attitudes about incidental findings suggests that in general they are viewed positively by the lay public and that the majority of people would wish to receive this information if it became available.

As with much of the information contained in this chapter, the approach to incidental findings is likely to evolve and may even have done so by the time you read these words. Investigating the latest ACMG recommendations is therefore recommended for the inexperienced clinician offering an NGS-based test.

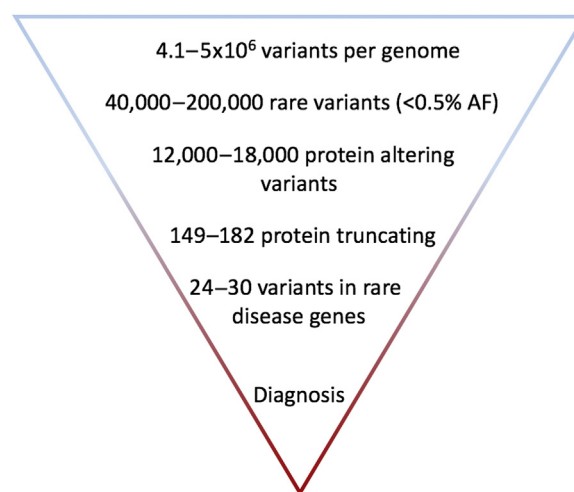
Patients sometimes turn down the offer of genetic testing, both for undiagnosed conditions and for conditions that run in their family. The reasons are numerous and varied as the patients making these decisions. The autonomy of the patient should be respected in all clinical decisions where competency is demonstrable, and there is no negative impact on society.

This raises a thorny issue of when to offer genetic testing in children who are too young to consent for themselves. The generally adopted approach in the United Kingdom is to test at the discretion of the parents for conditions that are likely to have their onset during childhood, but to refuse testing for adult onset conditions.

This presents no particular ethical boundary in performing NGS to reach a diagnosis for the patient in our given example. His developmental delay is an existing condition with onset in childhood, or even earlier.

Take again our example of an incidental finding of PSEN1—early-onset Alzheimer's. To this author's knowledge, no childhood onset cases have been recorded, and therefore testing would usually be delayed until a patient is old enough to ask for it themselves and consider the implications of diagnosis before proceeding.

The reporting of the ACMG minimal list is also therefore fraught in children. Many cancer predispositions do not manifest until adulthood; however, failing to report a pathogenic mutation would prevent the patient accessing either high-risk screening programs or prophylactic treatments that might otherwise have been beneficial when the time comes. Although there is still no firm consensus on the matter, generally if management options



**FIGURE 4.3** Variants per genome found in 1000 genomes project. Genomes from African populations generally have numbers at the higher end of the ranges stated, in keeping with the “out of Africa” theory of human migration. Different ethnicities fall at different points within each range. AF, Allele frequency.

exist for the condition then disclosure of IFs is ethically permissible even in a condition that will not manifest until adulthood.

In summary the nature of the data generated by NGS and its capacity for generating nonhypothesis-driven diagnoses present unique ethical challenges. It is the responsibility of any clinician offering NGS as a diagnostic service to be aware of these challenges and provide the appropriate guidance to their patients. The process of consent for NGS and specific consent forms should be used, so that the consent can be recorded in the patient's medical records. Incidental findings of any type should be considered carefully and delivered to the patient in an appropriate setting, with follow-up plans in place (Fig. 4.3).

## 4.7 Bioinformatics

Bioinformatics can be loosely defined as computational programs that handle biological data. In recent times the term has become more associated with pipelines for the analysis of high-throughput genomic data.

Although bioinformatics is not usually part of a clinician's workload, it is important to appreciate the processing that takes place and the skill that goes into designing and managing these pipelines. It is also important to recognize that changes to the bioinformatic pipeline can result in different variants being identified.

The raw data produced by NGS is of little use without significant processing. A bioinformatic pipeline will use the data produced to generate a list of variants present in the sequence data compared to the reference genome being used. Sequence generation tends to be handled by the platform doing the sequencing. It takes the light pulses, colored fluorophore releases, or electrical currents generated by the flow cell and converts them into a sequence of bases and quality control information. This information is stored in a standardized, text-based format called a FASTQ file.

The raw sequence data contained in the FASTQ files then needs to be compared to a reference genome or exome, for differences to be identified. This begins with a process called alignment, in which the sequences are mapped to the reference genome.

There is a suite of tools available for sequence alignment, but perhaps the most widely used is the Burrows–Wheeler aligner (bwa) package. The choice of package is dependent on the characteristics of the sequence data—for example—bwa itself has different algorithms for the alignment of short sequences (<100 bp) and longer sequences (100–1 Mb). Third-generation sequencing data has different characteristics to second-generation data, so novel tools have emerged for handling of this data.

Once the reads have been aligned in a SAM file, converted to a binary format (a BAM file) and indexed, so that they can be easily searched by downstream tools, variant calling can take place. A variant caller, such as Bambino, identifies places in which the aligned sequences are different from the

The output of variant calling is stored in a variant call file or binary call file. Each variant file will contain many thousands of variants, with quality scores for each derived from the aggregate quality of the reads from the aligned FASTQ file, the read depth at the variant and the number of strands in which the variant was identified.

Based on data from the 1000 genomes project, it is estimated that genome sequenced will contain 4–5 million variations from the current build of the reference genome (around 0.15% of the genome) (Fig. 4.3). Obviously, analyzing each individual variant would take a clinician several lifetimes, so the variants must be filtered to generate a useful report.

Variant filtering involves several steps. The first is to exclude common variants, with common being defined as present in >0.5% of reference alleles. At present, intronic and intragenic variants are excluded as interpretation is not possible with our current understanding of genomics. It is possible, perhaps even likely, that these variants could interfere with regulatory regions and cause disease; but at present, the state of our knowledge does not allow use for clinical diagnosis.

The next is to exclude obviously benign variants, such as synonymous variants in protein-coding regions. These can be further filtered to identify variants in possible disease genes.

An *in silico* panel (ISP) can also be applied to the filtering process. ISPs are lists of genes known to be associated with a particular presenting complaint. The 100kGP, the successor to the 1000 genomes project, maintains a curated list of ISPs on the Panel App website.

ISPs can reduce the number of VUS in genes not relevant to the condition, as well as reducing the chance of identifying incidental findings. The variants in genes excluded from the panel are not analyzed further or reported.

As the variants are filtered, they will also be annotated with population frequencies and *in silico* predictions of their impact. For further information on these aspects, please refer to the next section on clinical interpretation.

The result of this process is a short list of protein altering or truncating variants that have been identified in the exons or splice sites of genes known to be associated with the presenting condition. Unless there is an obviously pathogenic variant these candidate variants are then taken forward for clinical interpretation, usually in a multidisciplinary team setting.

## 4.8 Clinical interpretation of variants

Upon receiving a report from NGS, a clinician may be faced with one or more variants that could be responsible for the patient's presenting features. Interpretation of these variants is a complex and sometimes controversial process. In this section, we will review the practice of variant interpretation and the resources that are available to assist with this process. During the course of the review, we will discuss the ACMG guidelines on variant interpretation, as well as some of the criticisms of these guidelines.

The ACMG guidelines for variant interpretation are an attempt to standardize the classification of variants by grouping them into one of five categories: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign.

Each variant is scored according to the following criteria:

1. Population data
2. Computational and predictive data
3. Functional data
4. Segregation data
5. De novo status
6. Allelic data
7. Other databases
8. Other data

The scoring system is complex, and there is little sense in copying the entire system from the paper into this chapter. Instead, a condensed overview of each criterion is offered. It is always recommended to refer directly to the latest version of the ACMG guidelines when interpreting variants.

### 4.8.1 Population data

Several databases exist, cataloging variants from various studies, which have used WES and WGS. These include ExAC, GNOMAD, dbSNP, dbVar, 1000 genomes project, and EVS. Bioinformatic pipelines will often annotate candidate variants with a comment about whether the variant is present in population databases.

Absence of the variant from a database, or presence in affected patients both count toward the classification of a variant as pathogenic. A mean allele frequency too high to be consistent with the disease (i.e., a MAF of 1 in 1000 for a condition that affects 1 in 50,000 people) is considered to be evidence that a variant is benign.

### 4.8.2 Computational and predictive data

Predictive evidence is either based on previously published disease associations, or on in silico modeling tools that predict the functional consequence of a variant on the protein. Evidence taken from previously established associations is weighed more heavily than computational predictions.

A loss-of-function (LOF) mutation, which is a nonsense mutation or a truncating frameshift mutation, in a gene where LOF has been *previously* described as a mechanism of disease is very strong evidence of pathogenicity. Some genetic disorders are caused by gain in toxic function rather than LOF, so an understanding of the pathology underlying the disorder associated with the gene is important.

Similarly, a missense mutation that leads to an amino acid change that has been previously associated with the condition can be taken as evidence of pathogenicity.

The Online Mendelian Inheritance in Man (OMIM) database has extensive and up-to-date information about disease associations and details of some previously reported mutations. It is a reliable resource for finding out whether the variant identified has been previously reported or similar variants have been known previously cause the condition.



ClinVar is a curated database of single-nucleotide polymorphisms (SNPs), with clinical annotation. The level of evidence supporting benignity or pathogenicity is also rated using a star system, for ease of interpretation.

PubMed, the database of published medical journals, can also be checked to find evidence that may not yet have been included in these curated databases. Either by searching for the gene by name, or the name of the associated condition, up-to-date information can be sought.

For unique variants that have not been previously reported, in silico predictive tools exist to estimate the effect of the variant on the protein produced. They generally use a combination of evolutionary conservation and severity of chemical change to make their predictions.

If an amino acid has been unchanged in a protein that has been present across millions of years of evolution then it is reckoned that the amino acid in question is vital to protein function. Missense changes affecting this amino acid have not resulted in viable offspring, and therefore variants at this locus have been “bred out.” Similarly, if an amino acid shows a high degree of variation between different species then it can be reckoned that variation is more tolerated at this locus and less likely to cause disease.

Not all amino acid changes have equal capacity to damage a protein. Chemical properties such as charge, hydrophobicity, and relative pH of each amino acid exist on a scale. If a missense mutation causes a change in an amino acid that completely changes the properties of that locus on the protein then it is more likely to affect the protein’s function than a switch to a chemically similar residue.

These in silico tools synthesize this information in order to make a prediction as to the functional impact of a variant on a protein. They are known to be more sensitive than specific and false positive calls are relatively common. The guidelines therefore suggest that *multiple* lines of computational evidence are required to count toward the classification of a variant. Even when there are multiple lines of evidence in agreement, this is only considered weak evidence.

### 4.8.3 Functional data

Functional data can be considered either established laboratory tests that provide supporting evidence for the negative impact of a mutation or inferred functional data from previous publications.

Inferred data includes a missense mutation in a gene with a low rate of missense variants or a high rate of pathogenic missense variants. This data can be found on population databases such as those described earlier.

Both ExAC and DECIPHER provide a haploinsufficiency index score. Lower scores (closer to 0%) indicate that a gene is more likely to exhibit haploinsufficiency and that losing one allele could cause disease.

They also provide a loss intolerance score—indicating whether LOF has been reported as a pathogenic mechanism in the gene. Scores of <50% indicate recessive pathogenicity, whereas scores of <10% indicate dominant pathogenicity for LOF mutations.

ExAC also provides a Z score for the number of missense mutations seen in a gene compared to the number expected. Genes with far fewer missense mutations than expected are considered likely to be intolerant of missense mutations.

If a mutation occurs in a well-studied mutational hotspot or functional domain, such as the binding site of a transcription factor, then again, this can be considered evidence for pathogenicity.

The most significant evidence comes from validating the variant with established functional laboratory tests. This can take the form of enzyme assays, RNA sequencing to examine transcription, biopsy of effected tissue, or any other test where a result is confirmatory of a diagnosis. Arranging these tests requires a fair deal of clinical acumen and can be a time-consuming process, for example, a patient with VUS in a gene associated with neuropathy may need neurological assessment and referral for nerve conduction studies to provide functional support for the variant classification.

Unsurprisingly, functional studies that do *not* show the expected impact associated with the variant being assessed can be considered strong evidence against pathogenicity.

### 4.8.4 De novo status and segregation data

In genes that cause disease in an autosomal-dominant fashion, it is important to know if the variant was inherited from a healthy parent, or if it has occurred de novo. De novo mutations usually arise from changes in the genome of the sperm from which a child is conceived. Trio sequencing will automatically filter to exclude inherited variants in these genes.

If proband-only sequencing has taken place (and both parents are available) then testing for the identified variant can take place once it has been identified.

In cases where there is a family history of the condition, segregation analysis should be performed to discover whether the variant is always present in affected family members (or obligate female carriers in the case of X-linked conditions). The more generations in which a variant segregates with disease, the stronger the case for pathogenicity. If the variant is found in any unaffected family member for a condition that is expected to be fully penetrant, then this is strong evidence against pathogenicity.

#### 4.8.5 Allelic data

Allelic data is especially important in the classification of variants that are associated with recessive conditions. If there are two potentially pathogenic variants that occur on the same allele then they are not the cause of the patient's disease. Recessive variants must exist in *trans* in order to cause a disease phenotype.

Similarly, if two dominant variants exist in *cis*, which are on the same allele, then it is very likely that only one is pathogenic. Also, if two dominant variants exist in *trans*, it is likely only one is pathogenic as loss of both copies of a gene that exhibits haploinsufficiency often confers a catastrophic phenotype.

#### 4.8.6 Other databases

Besides the large population databases mentioned, there are disorder-specific databases for many of the more common genetic diseases, run by centers or research groups with particular expertise in these conditions. Variants are sometimes found on these databases that have not reached the larger databases, and this can be considered toward variant classification

#### 4.8.7 Other data

If the patient's phenotype or family history is particularly specific for the gene in which the variant occurs then this can be used as supporting evidence for pathogenicity.

#### 4.8.8 Criticism of American College of Medical Genetics and Genomics guidelines

It should be recognized that the ACMG guidelines for variant interpretation are a first attempt by genetic health professionals to create a predictable and reproducible system for the classification of variants discovered by NGS. Studies examining the use of these guidelines suggest that although the terminology used has become more standardized, the classifications reached are not. One study found that concordance was as low as 20% between different professionals using the guidelines to classify the same variants.

Another issue that has been identified is that functional research into the effects of newly identified mutations occurs at a much slower rate than identification of new mutations. High-throughput functional cell biology does not yet exist, so the use of clinical information to interpret variants retains a position of primacy in variant interpretation. Therefore there are also a large number of genes for which no established functional laboratory investigation exists.

In light of this situation, it has been suggested that clinical criteria could be given a greater weighting in a revised scoring system. The SHERLOC (semiquantitative, hierarchical evidence-based rules for locus interpretation) system is based on ACMG guidelines, which takes both of these issues into account.

This is a recent development, and it remains to be demonstrated if this alternative approach leads to a greater consensus between the clinicians engaged in variant classification.

#### 4.8.9 Summary: potential future developments in clinical genomics

The main focus of this chapter has been on the interpretation of Mendelian genetic variants that confer a disease phenotype when present in homozygous, heterozygous, or compound heterozygous state. One area of considerable interests at present is the contribution of low-frequency SNPs toward genetic risk. Large consortia have published lists of SNPs associated with small increases in the polygenic risk for several common conditions including type 2 diabetes mellitus, breast cancer, and schizophrenia.



Both research into polygenic risk alleles and identification of disease genes that have not yet been identified require large research consortia that can sequence and analyze on an industrial scale.

The most prominent project of this type currently running in the United Kingdom is the 100kGP. The 100kGP is an ambitious project with the aim of embedding WGS use in the health service. As the name implies, the project aims to sequence 100,000 genomes; an order of magnitude greater than have been sequenced in total so far.

The bandwidth for the generation of genomic data is in danger of outpacing the storage requirements necessary to store that data. Given the confidential nature of information arising from diagnostic tests and the increasing threat of cybercrime threatening the informatics infrastructure of the health service, the issue of adequate secure data storage will present structural challenges that must be overcome if WGS is to be utilized on a wider scale.

Data storage also needs to proceed in a manner that allows access to the data files for further analysis should they be required. In the first instance, this may be desirable if a patient develops a new condition that may have a genetic predisposition (breast/ovarian cancer at a young age, for example). Stored genomic data can negate the need for further testing.

Recent publications have highlighted the benefit of reanalysis of stored NGS data files. As many as 10% of patients with no pathological variants identified at the time of sequencing could receive a diagnosis if the data is reanalyzed after several years. Advances in bioinformatics contribute to this extra pick-up rate—so does the identification of new gene disease associations.

The capacity for the reanalysis of existing data in many clinical settings is limited; however, in certain cases, it may be warranted. It is certainly less expensive than resequencing the patient's genome using new technology.

Other developments that could increase the utility of clinical NGS in coming years include whole-transcriptome sequencing and methylome sequencing. Transcriptome sequencing examines the relative abundance of RNA to determine whether a mutation in a specific gene has an impact on the quantitative expression of other genes. Many genes regulate and coregulate the expression of other genes, and transcriptome sequencing has the potential to allow network analysis of the impact of variants in these regulatory genes.

Methylome sequencing examines genome-wide DNA methylation. As methylation is known to regulate the transcription of genes, changes in the methylation state of individuals with normal WGS results could account for diagnoses that could not otherwise be made.

Synthesis of information from phenotype, genome, methylome, transcriptome, proteome, and metabolome sequencing has the potential to provide deep insight into the biological systems of an individual as a whole. Although such panomic analysis is still in its infancy, it is certainly possible that by the next edition of this book, this chapter will need to be expanded considerably to include new developments in these fields.



## S E C T I O N I I

---

# Molecular medicine in clinical practice



# Molecular basis of obesity disorders

*Lotte Kleinendorst and Mieke M. van Haelst*

Department of Clinical Genetics, Amsterdam University Medical Centers, Amsterdam, The Netherlands

## 5.1 Introduction

Obesity is an accumulation of body fat caused by an imbalance of energy intake and energy output. According to the World Health Organization (WHO), adult obesity is defined as a body mass index (BMI) greater than or equal to 30 kg/m<sup>2</sup>. Childhood overweight or obesity is evaluated by weight-for-height for children younger than 5 years old or BMI-for-age for children older than 5 years old. BMI is often used as a simple tool to determine healthy weight in the doctor's office (Table 5.1). Unfortunately, BMI is not the ideal parameter for measuring healthy weight or fat mass. Since BMI does not measure body fat, a very muscular person with low body fat could appear obese when one only looks at the BMI. Measurement of abdominal fat is therefore a better tool for assessing healthy weight, as it also has a stronger association with metabolic health outcomes such as type II diabetes than BMI.

### 5.1.1 Consequences of obesity

Obesity is a worldwide epidemic. The WHO's fact sheet on obesity and overweight reported that 650 million adults were obese in 2016. Furthermore, 381 million children and adolescents were overweight or obese [1]. The prevalence of obesity is increasing at a fast pace, and it is estimated that one-fifth of the world population will be obese by 2025 [2].

The obesity epidemic is extremely worrisome because of the various severe health risks that affect all organ systems. Important diseases in which obesity is a causative factor are type 2 diabetes, cardiovascular or cerebrovascular diseases, and several types of malignancies, such as breast cancer or colon cancer. It is estimated that obesity is the cause of cancer in approximately 10% of the cancer deaths of American nonsmokers [3]. An increased BMI is also a major risk factor for medical problems, such as joint problems (gout and osteoarthritis), cholelithiasis, and menstrual disorders, which disturb daily life. Moreover, the psychological effects of obesity should not be overlooked. It is known that obese children are more likely to develop psychological or psychiatric problems than children with a healthy weight [4].

### 5.1.2 Obesity: nature or nurture

Changing environmental obesogenic factors (the availability of cheap high-sugar or high-fat food and the decrease in physical activity) are the main cause of the increasing prevalence of obesity. However, since every person does not become obese in this obesogenic environment, it is likely that there are other factors that play a role in the development of obesity. It is still unclear what the exact contribution of genetic factors in the development of obesity is; numerous twin studies were performed to estimate the heritability of obesity. A metaanalysis of more than 2500 worldwide twin studies published between 1950 and 2012 showed a heritability of 72.6% for the trait "weight maintenance" [5]. Similar results were found in adoption studies [6]. This suggests that weight is a highly heritable trait, meaning that the variation in weight in a population is highly attributable to genetic factors.

**TABLE 5.1** Body mass index (BMI) classification according to the World Health Organization International Classification of adult underweight, overweight, and obesity.

Classification	BMI (kg/m <sup>2</sup> )
Underweight	<18.5
Normal	18.5–24.99
Overweight	≥ 25.00
Class I obesity	30.00–34.99
Class II obesity	35.00–39.99
Class III obesity	≥ 40.00

### 5.1.3 Genetic obesity

Thus far a small percentage of patients with obesity have a proven genetic defect as the main cause of their obesity. This group of patients is often divided into patients with monogenic syndromic and monogenic nonsyndromic obesity. Syndromic obesity describes the patient that is not only obese but also has intellectual disability, congenital malformations, and/or dysmorphic features. The patients with nonsyndromic obesity mostly have one manifestation of their monogenic disorder: obesity. In this chapter, both types of genetic obesity are discussed.

## 5.2 Clinical cases

### 5.2.1 Case 1

Patient A, a 14-year-old boy with early-onset obesity, severe hyperphagia and a large head circumference.

Patient A is a 14-year-old boy with a height of 181 cm (+1.5 SD for his age and gender group) and weight of 137.9 kg. His BMI at time of referral was 42.1 kg/m<sup>2</sup>. The pediatrician referred him to the Department of Clinical Genetics because of his severe hyperphagia. Endocrine causes of his obesity were already excluded by the pediatrician.

#### 5.2.1.1 Patient history

Patient A was following regular education and had no congenital anomalies. He was born at term after an uneventful pregnancy, with a birth weight of 4000 g. There was no developmental delay, and he was a healthy toddler. Patient A's rapid weight gain started at the age of 3. The referral letter mentioned that the suspicion of a genetic cause was initially low because there were psychosocial factors that could explain the obesity. He followed regular education, and there were no congenital anomalies or major illnesses in his medical history.

#### 5.2.1.2 Physical examination

During the clinical evaluation at the clinical genetics department, he was found to have a large head circumference of 60.5 cm (> +2.5 SD). There were no dysmorphic features that could be indicative of syndromic obesity.

#### 5.2.1.3 Family history

Patient A (Fig. 5.1) is the second child of nonconsanguineous Dutch parents. His older sister also has obesity. Her obesity started in the first year of life. She underwent bariatric surgery (gastric sleeve) at the age of 18 years. Her height is 168 cm. Before bariatric surgery she weighed 167 kg, with a BMI of 57.8 kg/m<sup>2</sup>. Her current weight is 158 kg. Their mother is 159 cm tall and weighs 114 kg. She underwent gastric banding, resulting in a total weight loss of 60 kg. Because of complications of the gastric band, 15 years later the band was removed and followed by a gastric sleeve operation. The mother's sister and mother also have obesity. The maternal grandmother underwent gastric banding with sufficient weight loss.





**FIGURE 5.1** Patient A with his sister and mother. Permission for the publication of this image was obtained from all three individuals. Source: Picture made by the Department of Medical Photography in the Amsterdam University Medical Centers, location AMC.

#### 5.2.1.4 Genetic diagnosis

Sanger sequencing of the melanocortin-4-receptor (*MC4R*) gene was performed to find a genetic diagnosis for patient A's obesity. The test showed a heterozygous pathogenic mutation: c.105C > A, p. Tyr35\*. This mutation is a known pathogenic mutation associated with early-onset obesity. Segregation analysis in the family showed that patient A's mother and sister also have the pathogenic mutation in *MC4R*.

### 5.2.2 Case 2

Patient B, a girl with early-onset obesity and severe hyperphagia. She (Fig. 5.2) was born at 34 weeks of gestation with a birth weight of almost +2 SD above the mean for her gender and duration of the pregnancy. A few weeks after birth, her parents began to notice severe hyperphagia. She became obese at the age of 2 months. The physician in the child health-care center referred the patient to a pediatrician, where dietary advice was given. At the age of 6 months, patient B was too heavy to fit in a normal baby carriage. Since the nutritional guidelines were not helpful, she was referred to an academic center for evaluation of pediatric obesity.

Her leptin serum level was normally elevated for her weight (57.7 µg/L—reference range 15–100 µg/L). Indirect calorimetry showed a 24% lower basal energy expenditure than normal for her weight.

#### 5.2.2.1 Genetic diagnosis

Elevated leptin levels excluded leptin deficiency. Initial genetic analysis of the *MC4R* gene showed no mutations. A year later the multigene panel test for obesity became available. This test showed two mutations in the leptin receptor (*LEPR*) gene: c.1985T > C p.(Leu662Ser) and c.2168C > T p.(Ser723Phe). The mutations were confirmed by Sanger sequencing, and segregation analysis showed that the parents are both carrier for one of the mutations. The variants were classified as variants of unknown clinical significance, since they have never been described before in other patients. Because the mutations affect highly conserved regions of the gene, it is likely that they have an important role in the receptor's functioning.

### 5.2.3 Case 3

Patient C is a girl with developmental delay, hyperphagia, and sleeping problems. She (Fig. 5.3) is the second child of nonconsanguineous parents. She was born at term after an uneventful pregnancy, with a normal birth



**FIGURE 5.2** Picture of patient B, 2 years old. Permission for the publication of this image was obtained from both parents. Source: *Photograph provided by the parents of patient B.*



**FIGURE 5.3** Picture of the patient at age 8. The patient does not have a dysmorphic facial appearance. She does have a relatively large occipitofrontal circumference. Permission from both parents for the publication of this image was obtained. Source: *The picture was provided by the patient's parents.*

weight. She demanded remarkably more feeding than her sibling. She also had a speech delay and difficulties trying to stand up. She showed signs of sensory sensitivity: she started to scream when exposed to bright lights or noisy crowds. Because of her sleeping problems and frequent infections at the age of 2.5 years, evaluation was requested at an outpatient clinic for children with developmental delay.

### 5.2.3.1 Genetic diagnosis

Because of the combination of developmental delay and a large head circumference, a SNP (single-nucleotide polymorphism) array analysis was performed. This showed the classical deletion at chromosome 16p11.2. Parental analysis showed that the deletion occurred *de novo* in the patient.

### 5.2.3.2 Follow-up

Patient C started to gain weight from the age of 7. She had hyperphagia and needed to be controlled in her food intake by her parents. At the age of 8 years, her height was 138 cm (+0.3 SD), and her weight was 45 kg with a BMI of 23.7 kg/m<sup>2</sup> (+2.8 SD). Her occipitofrontal circumference is 55 cm (+1.8 SD). Her IQ was tested at 81 and follows special education. She had autistic features, but no formal diagnosis of autism.

## 5.3 Molecular systems underpinning the clinical scenario

### 5.3.1 Case 1. Melanocortin-4-receptor gene mutations

#### Case 1. MC4R (Melanocortin 4 receptor)

Location: 18q21.32

Inheritance: autosomal recessive or autosomal dominant

Key feature: hyperphagia leading to early-onset obesity

Mutations in *MC4R* are the most common genetic cause of obesity without intellectual disability. In the population with severe childhood obesity, its prevalence is around 6% [7]. *MC4R* deficiency can be inherited as an autosomal-recessive condition as well as an autosomal-dominant condition. A single (heterozygous) pathogenic mutation in *MC4R* causes severe hyperphagia that leads to obesity. People with homozygous or compound heterozygous *MC4R* mutations (two mutations, one inherited from each parent) have a more severe hyperphagic and thus obese phenotype. Moreover, they can have hyperinsulinemia and reduced bone density. *MC4R*-deficient patients have an accelerated linear growth in childhood leading to an increased final height and a large head circumference. This might be caused by hyperinsulinemia or by effects on growth hormone secretion [8]. *MC4R* plays a crucial role in the leptin–melanocortin pathway, the hypothalamic pathway that controls food intake and energy expenditure (see Fig. 5.5). There are two neuropeptides that can bind to the *MC4R*:  $\alpha$ -melanocyte-stimulating hormone ( $\alpha$ -MSH) and agouti-related protein (AGRP). AGRP and  $\alpha$ -MSH compete to bind to receptor. The binding of  $\alpha$ -MSH leads to activation of the receptor, causing an anorexigenic signal (a signal that leads to a decrease in food intake). AGRP causes deactivation of the *MC4R*, causing an orexigenic signal (leading to increased food intake).

### 5.3.2 Case 2. Leptin receptor deficiency

#### Case 2. LEPR (Leptin receptor)

Location: 1p31.3

Inheritance: autosomal recessive or autosomal dominant

Key features: hyperphagia leading to early-onset obesity, hypopituitarism in patients with homozygous or compound heterozygous mutations.

Patient B was diagnosed with a compound heterozygous *LEPR* deficiency. Like the *MC4R*, the *LEPR* plays an important role in the leptin–melanocortin pathway (Fig. 5.5). Leptin is a hormone that is mainly produced and secreted by adipose tissue. The amount of leptin in blood increases, when adipocytes increase in size. This is why almost all people with adiposity have elevated leptin levels. When leptin binds to the *LEPR*, this activates the hypothalamic processes that affect the energy balance by inhibiting appetite. In the case of *LEPR* deficiency, this pathway is disrupted, causing hyperphagia and early-onset severe obesity with various forms of hypopituitarism: hypogonadotropic hypogonadism, growth hormone deficiency, and/or hypothyroidism. Affected patients often show a reduced final height because the pubertal growth spurt can be absent. *MC4R* agonists could be a promising therapy for weight loss in *LEPR*-deficient patients, since it bypasses the defective *LEPR* in the leptin–melanocortin pathway (see Section 5.5). There is evidence that patients with a heterozygous pathogenic mutation in *LEPR* are at risk for obesity, but they do not exhibit the hormonal problems that are present in patients with homozygous or compound heterozygous mutations.

### 5.3.3 Case 3. 16p11.2 deletion

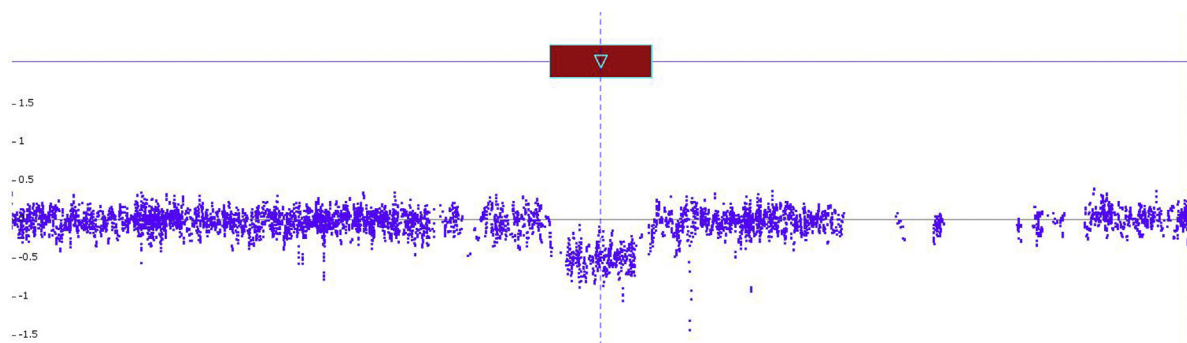
#### Case 3. 16p11.2 deletion

Location: approximate position of 29.6–30.2 Mb in reference genome GRCh37/hg19, including around 25 genes

Inheritance: autosomal dominant, but mostly de novo  
Key features: developmental delay, autism, and macrocephaly

Patient C was diagnosed with a 16p11.2 deletion, a 600 kb loss of the short arm of chromosome 16. The deletion was identified using SNP-array analysis (Fig. 5.4).

There is a large variation in symptoms and severity of symptoms among patients with a 16p11.2 deletion. Most patients have mild intellectual disability. The delay in language development is often more severe than the motoric development. Behavioral problems are frequently described, including autism and attention deficit hyperactivity disorder (ADHD). Around 20% of the patients suffer from epileptic seizures. Many have a relatively large head circumference. Moreover, these patients can suffer from constipation, sleeping problems, and scoliosis. The obesity phenotype in patients with a 16p11.2 deletion occurs later in childhood than in patients with mutations in the leptin–melanocortin pathway genes. There is a 43-fold increased risk to become morbidly obese for 16p11.2 deletion patients. Interestingly, individuals with a duplication at 16p11.2 have an increased risk of being underweight. This “mirror phenotype” does not only involve the obesity phenotype in 16p11.2 deletion and duplication patients, but it is also the case for head circumference. The deletion carriers often have macrocephaly, whereas the duplication carriers have microcephaly [9]. The 16p11.2 region includes around 25 genes. One gene in the 16p11.2 deletion that might be responsible for the obesity phenotype is *SH2B1*. Patients with a deletion including only *SH2B1* or patients with mutations in *SH2B1* are more likely to overeat and have rapid weight gain and insulin resistance [10]. *SH2B1* plays a supporting role in the leptin–melanocortin pathway, probably because it enhances the leptin sensitivity of the hypothalamus [11].



**FIGURE 5.4** The typical image of a 16p11.2 deletion identified with SNP-array analysis. The red arrow bar marks the region of the deletion. SNP, Single-nucleotide polymorphism. Source: Figure created by Elles Boon (Department of Genetics Amsterdam UMC) for Lotte Kleinendorst.



## 5.4 Molecular biology and pathophysiology of obesity

The human body consists of various different tissues including bone, muscle, nervous tissue, body fluids, and fat. Adipose tissue in the human body is located under the skin (subcutaneous fat), surrounding the organs (visceral fat), and in muscles and bone marrow. The typical human body is made up of 25%–30% adipose tissue in normal-weight women and 20% in men. Adipose tissue can be divided into white and brown adipose tissue. White adipose tissue functions as a storage facility for energy, while brown adipose tissue actually uses the stored energy, mainly to produce heat for body temperature regulation. Obesity results from an excessive amount of white adipose tissue. Not only the quantity of body fat is important for health effects, the location of the adipose tissue also influences the risk for heart disease, diabetes, and death. Central obesity, an excess of abdominal fat, leads to more health complications than fat located around the hips or the buttocks.

### 5.4.1 Fat storage

The two most important energy sources of the body are fatty acids and glucose. Glucose is stored in the form of glycogen. Glycogen yields about 4 kcal/g when metabolized. But since glycogen binds to large amounts of water, the energetic value of glycogen in the body is only 1–2 kcal/g. Fatty acids are stored in the form of triglycerides, and they have a much higher energetic value: 9 kcal/g. Fat storage in the body is a crucial biological process that makes us able to survive periods of low food supply. Adipocytes, or fat cells, store fat that is either consumed directly or metabolized from protein or sugar in the liver. During physical activity or fasting, the energy from adipose tissue can be released into the blood in the form of glycerol and fatty acids. These processes are influenced by the hormones insulin and glucagon among others.

### 5.4.2 The big two

Obesity is caused by an imbalance between energy intake and energy expenditure. The main players in this imbalance are an unhealthy diet and decreased physical activity, phrased as “the big two” [12]. The authors postulate, however, that there are at least 10 other putative additional explanations for the increasing obesity prevalence, varying from assortative mating (individuals with obesity who might tend to procreate with people who are obese as well) to chemicals in our food chain that can disrupt the endocrine system.

### 5.4.3 Nongenetic causes of obesity

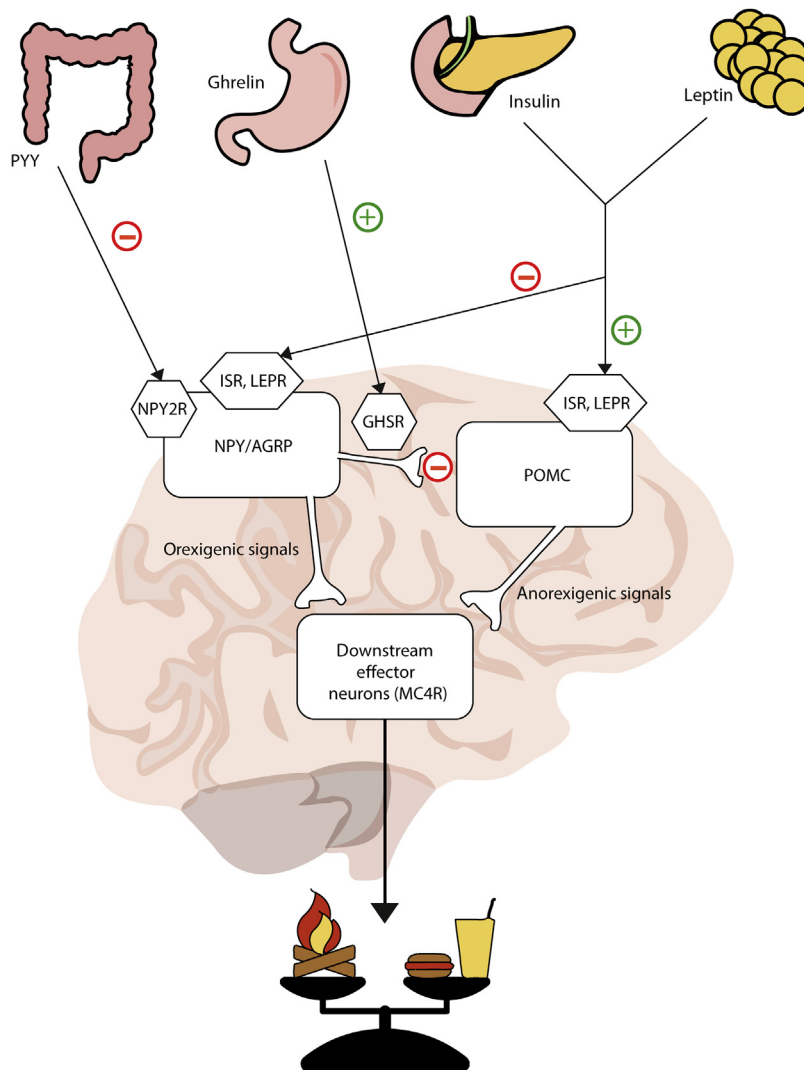
Besides an unhealthy diet and low physical activity, other lifestyle factors contribute to the development of obesity, such as chronic stress, sleep deprivation, pain, and crash diets. They probably influence weight through elevated cortisol levels. Glucocorticoids can increase appetite for high-energy foods and induce abdominal obesity [13]. This is also the case for patients with Cushing’s disease or inflammatory diseases for which glucocorticoids such as prednisone are used. Medication other than glucocorticoids can also be an important factor in weight increase, either by increasing appetite or by slowing down activity or metabolic processes. Antidepressants, antiepileptic medication, and insulin are well-known obesogenic drugs. Energy intake and expenditure are controlled by a complex neural system of which the most crucial processes take place in the hypothalamus. Therefore it is not surprising that severe hypothalamic damage from tumors, head injuries, or infections can also cause hyperphagia and obesity. These and other nongenetic causes of obesity are not further discussed elsewhere in this book.

### 5.4.4 The energy balance

Sensory information from all over the body reaches the brain through hormones or sensory nerves. Everyone knows how the mouth waters when you see, smell, or even think of a nice meal. After consuming this meal, the gastrointestinal tract is also communicating with the brain. The stomach, for example, is not only able to detect food by fibers of the vagal nerve in the mucosa, but it can also estimate the volume of food by vagal fibers in the gastric muscles that respond to stretching. The stomach’s mucosa also produces the hormones leptin and ghrelin. Ghrelin is secreted when the stomach is empty. When the stomach is full,

the stomach wall stretches and secretion of ghrelin stops. Ghrelin is a “hunger hormone” that stimulates a person’s appetite. Gastric leptin secretion occurs after food intake and leads to short-term satiety and therefore to a decreased food intake.

When nutrients come into contact with the gut, hormones such as cholecystokinin, glucose-dependent insulinotropic polypeptide, and neuropeptide YY (PYY) will be secreted and function as the signals between the gut and the brain. The release of these hormones is location- and calorie-dependent. PYY, for example, reduces appetite after a meal. Moreover, when glucose levels rise in the blood, the pancreas releases amylin and insulin, leading to appetite reduction via the hypothalamus (Fig. 5.5). Leptin is not only secreted in the stomach’s mucosa, it is mainly synthesized and secreted by adipose tissue. Leptin works through the LEPR that is widely expressed in the hypothalamus. Normally, leptin binds to LEPRs that are located on the hypothalamic neurons that produce proopiomelanocortin (POMC) and the neurons that produce AGRP. When the LEPR is activated, it stimulates the expression of POMC. POMC is cleaved into the melanocortins: MSH and into adrenocorticotrophic hormone (ACTH). MSH activates the melanocortin-4 receptor, leading to an anorexigenic signal, and thus to a lower food intake. This signaling pathway is called the leptin–melanocortin pathway. Another process that happens when the LEPR is activated is that the production of AGRP is inhibited. This process prevents the orexigenic effect of AGRP. When MC4R is bound by AGRP the receptor is deactivated. The AGRP pathway is also able to suppress the expression of POMC.



**FIGURE 5.5** A schematic overview of the different players in the leptin–melanocortin pathway leading to changes in energy expenditure and food intake through several downstream effector neurons including MC4R. *AGRP*, Agouti-related peptide; *GHSR*, growth hormone secretagogue receptor; *ISR*, insulin receptor; *LEPR*, leptin receptor; *MC4R*, melanocortin-4 receptor; *NPY*, neuropeptide Y; *NPY2R*, neuropeptide Y2 receptor; *POMC*, proopiomelanocortin; *PYY*, peptide YY. Source: Figure made by Lotte Kleinendorst.



### 5.4.5 Leptin resistance

Since people with obesity have more adipose tissue, they have higher leptin levels than normal-weight people. The high amount of leptin, however, appears not to lower appetite or increase thermogenesis as it does in normal-weight people. This process is named “leptin resistance”; leptin has lost its anorexigenic effect.

### 5.4.6 Not just the leptin–melanocortin pathway

The leptin–melanocortin pathway plays a crucial role in the balance of energy intake and expenditure, but other parts of the brain, such as the cortex and limbic regions, are important as well. External cues, such as seeing the logo of a fast food restaurant chain or hearing the jingle of the ice cream truck, interact with the emotional and cognitive parts of the brain. The reward system, inhibition control, and motivation to adhere to a diet are without a doubt critical players in the food intake system as well. Because most currently known genetic obesity syndromes affect the leptin–melanocortin pathway, higher cortical functions in relation to satiety and appetite are not further discussed in this chapter.

### 5.4.7 Genetic obesity: leptin–melanocortin pathway

Congenital leptin deficiency was first identified in a consanguineous family with two obese cousins from Pakistan. These patients had very low leptin levels in serum and suffered from early-onset morbid obesity. A homozygous frameshift mutation in *LEP* was identified in both children [14]. The first patient with *LEPR* deficiency was identified in a consanguineous family as well [15].

#### 5.4.7.1 Melanocortin-4 receptor

Melanocortin-4-receptor deficiency is the most common genetic form of obesity. In 1998 two independent research groups identified the first patients with *MC4R* deficiency at the same time, which led to a back-to-back publication of the articles “A frameshift mutation in *MC4R* associated with dominantly inherited human obesity” [16] and “A frameshift mutation in human *MC4R* is associated with a dominant form of obesity” [17] in *Nature Genetics*. *MC4R* is a G-protein-coupled receptor, a class of receptor that is crucial for signal transduction pathways, and therefore one of the most important drug targets in modern medicine. *MC4R* is a main player in the leptin–melanocortin pathway.

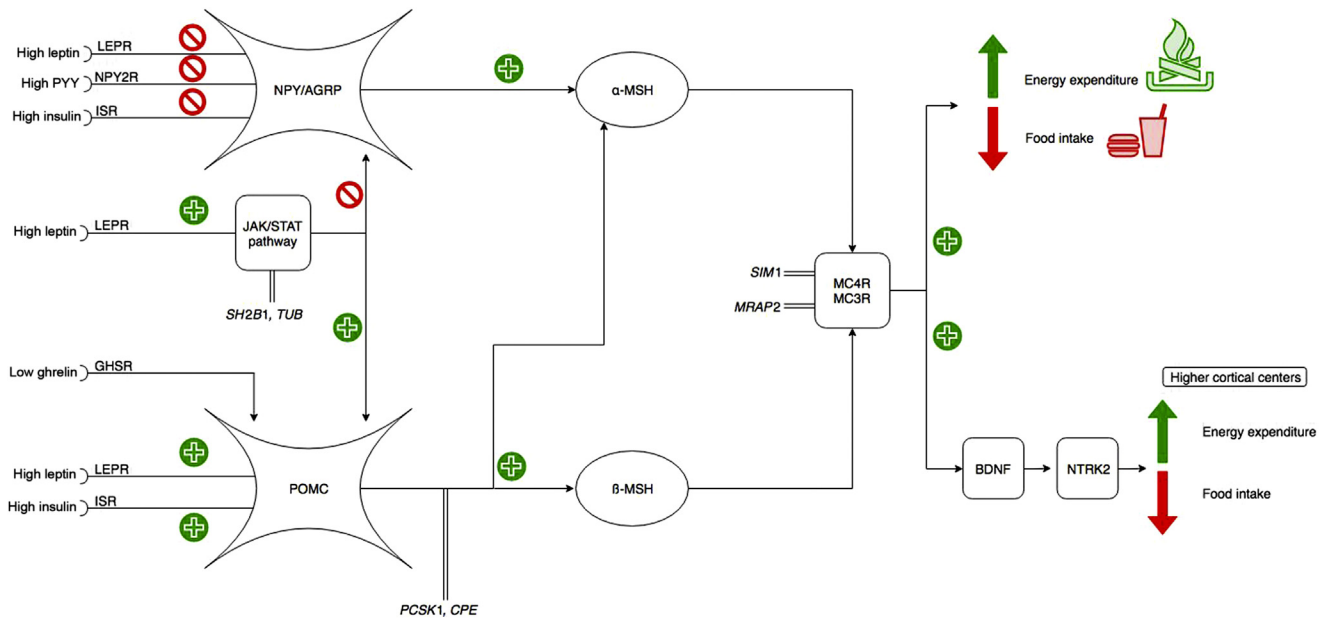
#### 5.4.7.2 Proopiomelanocortin

Autosomal-recessive mutations in the proopiomelanocortin (*POMC*) gene were first detected in 1998, in two nonrelated patients with morbid obesity and adrenal dysfunction [18]. Their striking phenotype included pale skin and red hair. This can be explained by the cleavage of the *POMC* protein into different peptides. One of these peptides,  $\alpha$ -MSH is needed for the production of skin and hair pigments. Deficiency of the second cleavage product, adrenocorticotrophic hormone (ACTH) results in adrenal dysfunction. Other examples of leptin–melanocortin pathway genes that are associated with obesity are *PCSK1* and *SIM1* (Fig. 5.6). *PCSK1* is important for the cleaving of *POMC* into the various peptides. *SIM1* is also expressed in neurons of the hypothalamus and is linked to *MC4R* signaling.

### 5.4.8 Monogenic obesity syndromes with intellectual disability

#### 5.4.8.1 Prader–Willi syndrome

Prader–Willi syndrome (PWS) is the best known syndromic obesity disorder and is caused by genetic changes within the Prader–Willi critical region on the long arm (q) of chromosome 15. The three main molecular problems that cause PWS are a paternal deletion of the Prader–Willi critical region, maternal uniparental disomy of chromosome 15, or an imprinting defect of the Prader–Willi critical region. The time line of the weight gain in PWS is very typical for the disease: from the neonatal period till the first year of life, a baby with PWS is hypotonic and has severe feeding difficulties due to poor suck and reduced appetite, for which tube feeding is almost always required. From the age of 1 year the children show an improved appetite and weight increase. The extreme hyperphagic phase that is typical for PWS often starts around the age of 8 [19].



**FIGURE 5.6** A more elaborate overview of the players in the leptin–melanocortin pathway, showing currently known genes associated with monogenic obesity. This figure shows the well-fed situation, in which leptin is secreted from adipose tissue. Leptin binds to the LEPR in the hypothalamus. Leptin binding to its receptor has three effects: (1) it inhibits NPY/AgRP production. (2) It promotes POMC production, which is then cleaved into the melanocortins  $\alpha$ - and  $\beta$ -MSH. This process takes place under influence of PCSK1 and CPE. Both  $\alpha$ - and  $\beta$ -MSH bind to the MC3R and the MC4R. This process is influenced by MRAP2 and by SIM1. MC4R activity leads to decreased food intake and increased energy expenditure. This energy effect also works through the higher cortical centers, via BDNF and NTRK2. (3) The third effect of leptin binding to LEPR is that it activates the JAK/STAT pathway. This pathway is influenced by SH2B1 and TUB and supports leptin's anorexigenic effect. BDNF, Brain-derived neurotrophic factor; CPE, carboxypeptidase E; JAK, Janus kinase; LEPR, leptin receptor; MC3R, melanocortin 3 receptor; MC4R, melanocortin-4 receptor; MRAP2, melanocortin 2 receptor accessory protein 2; NPY/AgRP, Neuropeptide Y/agouti-related protein; NTRK2, neurotrophic tyrosine kinase receptor 2; PCSK1, proprotein convertase, subtilisin/kexin type 1; POMC, proopiomelanocortin; SIM1, single-minded 1; STAT, signal transducer and activator of transcription; TUB, tubby bipartite transcription factor. Source: Figure made by Lotte Kleinendorst.

The exact mechanism of obesity in PWS is unclear, but a hypothalamic defect is most likely the cause of the lack of satiety in children with PWS. Defects in prohormone processing are currently studied as the main cause of hyperphagia in PWS. This could be caused by SNORD116 deficiency (SNORD116 is one of the genes in the Prader–Willi region), resulting in a decreased PCSK1 expression, [20]. Moreover, the neonatal phase with feeding problems could cause a long-term decreased caloric requirement in children with PWS. PWS patients do not have a growth hormone deficiency but, growth hormone therapy can help in the reduction of body fat and increase the amount of muscle leading to elevated energy expenditure.

#### 5.4.8.2 Bardet–Biedl syndrome

Another well-known genetic obesity syndrome is Bardet–Biedl syndrome (BBS), a disorder characterized by intellectual deficit, polydactyly, rod-cone dystrophy, and renal problems. There are more than 20 genes associated with BBS. Most of the patients, between 72% and 92%, have obesity [21]. Their weight gain often starts in the first year of life. There is more evidence available about the causes of obesity in BBS than in PWS. Leptin resistance is found in patients with BBS, similar to other patients with obesity. Since BBS is a disease of cilia dysfunction, it makes sense that cilia also play a role in the leptin–melanocortin pathway. Murine studies showed that the BBSome interacts with the LEPR [22]. Besides this, BBS patients show less physical activity than healthy controls [23]. Obesity in BBS is therefore probably caused by both increased caloric intake due to leptin resistance and lower energy expenditure.

### Learning points

- Genetic obesity syndromes can be divided into syndromes with and without developmental delay or intellectual disability.
- The most important group of genetic obesity syndromes without developmental delay is the syndromes caused by mutations in the leptin–melanocortin pathway. In these patients the hypothalamic satiety system is disrupted, leading to early-onset obesity due to hyperphagia.
- The most common obesity syndrome with intellectual disability is PWS. In contrast to the leptin–melanocortin pathway–related obesity disorders, the onset of obesity is later in childhood (> 5 years old).

## 5.5 Targeted molecular diagnosis and therapy

### 5.5.1 Targeted molecular diagnosis

#### 5.5.1.1 Leptin

Leptin concentration measurements in blood serum can help to diagnose leptin deficiency. Other types of genetic obesity, such as LEPR deficiency, cannot be effectively predicted by the concentration of leptin in the blood. But normal leptin levels are not even enough to prove that there are no pathogenic mutations in *LEP*. Mutations in the gene coding for the leptin hormone (*LEP*) can also cause inactive leptin, which is not able to activate its receptor [24]. So DNA diagnostics are needed when a genetic obesity syndrome is suspected.

#### 5.5.1.2 DNA diagnostics

At the moment the “Clinical Practice Guideline” of the Endocrine Society [25] suggests genetic testing in the following patient groups: (1) children with extreme early-onset obesity (before the age of 5 years), (2) patients that have clinical signs of a genetic obesity syndrome, and (3) patients with a family history of extreme obesity. Traditionally, the molecular diagnosis of genetic obesity syndromes has been based on time-consuming and labor-intensive gene-by-gene Sanger sequencing. Since some genetic obesity syndromes only present with early-onset obesity and hyperphagia without additional clinical features that suggest a genetic obesity diagnosis, a targeted next-generation sequencing obesity panel, whole-exome sequencing, or whole-genome sequencing should be considered. However, one should be aware that many of these tests are not (yet) able to identify every genetic obesity syndrome, for example, diagnoses of PWS and 16p11.2 deletions can be missed by using these techniques.

### 5.5.2 Prader–Willi syndrome

PWS is caused by genetic changes within the Prader–Willi critical region on the long arm (q) of chromosome 15 (paternal deletion, maternal uniparental disomy 15, or imprinting defect). Tests that can detect all three of these mechanisms are DNA methylation tests by Southern blot, methylation-specific-polymerase chain reaction (PCR), or methylation-specific multiplex ligation-dependent probe amplifications (methylation-specific MLPA).

### 5.5.3 Copy number variation

There are several copy number variations associated with obesity. The most common is a  $\pm 500$  kb deletion on chromosome 16p11.2. Copy number variations can be detected by array analysis such as array comparative genomic hybridization (CGH) or SNP array.

### 5.5.4 Therapy

#### 5.5.4.1 Lifestyle interventions

Since obesity is a problem of misbalance between energy intake and expenditure and it is often caused by exogenous factors, lifestyle adjustments are the first step in the treatment of obesity. A combined therapy of diet,

exercise, and cognitive behavioral therapy seems to be the most effective treatment. Unfortunately, many people find it difficult to adapt their lifestyle forever. Therefore people have been trying to find an easy solution for obesity for ages. One of the probably unsuccessful historical therapies for obesity was the cutaneous application of live leeches to suck the fat out of the blood (Fig. 5.7).

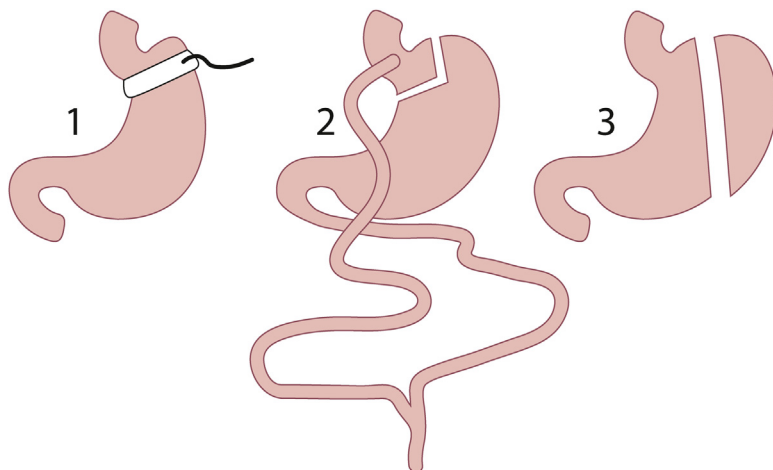
#### 5.5.4.2 Bariatric surgery

A more invasive and also much more effective therapy is bariatric surgery. The goal of bariatric surgery is to achieve weight gain by reducing the volume and uptake capacity of the stomach. There are strict indication rules for bariatric surgery in most countries. Patients have to be between 18 and 65 years old and have a BMI that is higher than  $40 \text{ kg/m}^2$ . Patients with a BMI between 35 and  $40 \text{ kg/m}^2$  can only qualify for bariatric surgery when they have one or more comorbidities, such as diabetes, obstructive sleep apnea, or cardiovascular disease. The three most typical procedures are gastric banding, Roux-en-Y gastric bypass, and sleeve gastrectomy (Fig. 5.8).

In gastric banding a fluid-filled ring is placed around the stomach. This restricts the passage of food through the upper part of the stomach, leading to increased satiety. The failure rate of gastric banding is relatively high, because patients with a gastric band still have to adhere to dietary instructions. It is, for example, still possible to drink large amounts of soft drinks with a gastric band. The laparoscopic Roux-en-Y gastric bypass is more effective than gastric banding and is currently the most performed bariatric intervention. This operation technique



**FIGURE 5.7** Leeches as treatment for obesity, unpleasant and unsuccessful. Source: Copyright free image, U.S. National Library of Medicine.



**FIGURE 5.8** The three most well-known bariatric procedures, from left to right: gastric banding, Roux-en-Y gastric bypass, and sleeve gastrectomy. Source: Figure made by Lotte Kleinendorst.

leads to weight loss because the stomach is converted into a small gastric pouch (around 30 mL, the volume of a single shot of espresso). Both the remnant of the stomach and the gastric pouch are connected to the intestine. The volume reduction of the stomach leads to important changes in neuroendocrine hormones that influence hunger and satiety. The secretion of ghrelin (the hunger hormone) by the stomach is reduced, while the secretion of PYY in the intestine is stimulated because food reaches the small bowel faster than normal. A rise in PYY will increase satiety. Sleeve gastrectomy is the second most performed weight-loss procedure. In this procedure a large part (around 60%–85%) of the stomach is removed, resulting in a tube-shaped stomach of around 100 mL. It is often performed in obese patients with a BMI above 50 kg/m<sup>2</sup>, when gastric bypass surgery is deemed risky.

Bariatric surgery is more effective in reaching weight gain than conservative treatment; the mean weight loss of all operation techniques is 61.2% [26]. Operations that lead to less than 50% of excess weight loss are considered “failed”; this happens in around 20% of the patients [27]. Why there is such a high variability in response to bariatric interventions is unknown, but genetic factors will probably play a role.

### 5.5.4.3 Bariatric surgery in genetic obesity

#### 5.5.4.3.1 Monogenic nonsyndromic obesity

Roux-en-Y gastric bypass is an effective procedure for MC4R deficiency. There is less experience with gastric sleeve resections in patients with MC4R mutations, but it seems effective as well [28]. This was unfortunately not the case for the sister in case 1, who is now only 9 kg lighter than before bariatric surgery. It is suggested that mutation carriers have more complications than patients without an MC4R mutation [29]. There is not much experience with bariatric surgery in patients with homozygous or compound heterozygous mutations in the LEPR gene. Case reports show both successful and unsuccessful outcomes of bariatric surgery in these patients [15,30].

#### 5.5.4.3.2 Prader–Willi syndrome

Sleeve gastrectomy appears to be an effective method for weight loss in children and adolescents with PWS. In PWS, patients aged between 4 and 18 years, sleeve gastrectomy was successfully performed in a 2015 study by Alqahtani et al. They found no complications or mortality after surgery in this group [31]. Bariatric surgery on patients with intellectual disability is not often performed, which makes this a controversial treatment for patients with PWS.

## 5.5.5 Medication

### 5.5.5.1 Nonpersonalized medication

There are few prescription drugs available for the treatment of “common obesity.” Many were taken off the market because of serious side effects. Orlistat, or tetrahydrolipstatin, is a drug that binds lipase in the digestive tract and therefore inhibits the hydrolysis of triacylglycerol into monoacylglycerol. This reduces the uptake of triglycerides in the digestive tract by 30%. With orlistat a weight loss of 2–4 kg can be achieved. However, a positive effect on morbidity and mortality was never proven, and many patients experience bowel problems, flatulence, and oily rectal leakage.

### 5.5.5.2 Personalized treatment

Currently, there is only one genetic obesity syndrome for which personalized treatment is scientifically proven and available. Leptin deficiency is treatable with daily subcutaneous injections of leptin. This treatment was first given in 1999 to a 9-year-old severely obese girl and led to a weight loss of 16.4 kg in a year [32]. There was hope that leptin treatment would also lead to weight loss in obese patients without leptin deficiency, but that was not the case. Many more personalized drugs were developed after the discovery of leptin treatment. A new and promising drug is setmelanotide, which is a melanocortin-4-receptor agonist that can be administered by subcutaneous injections. Phase 1 studies showed that it can lead to a higher resting energy metabolism in obese patients who were not tested for genetic causes of obesity [33]. Setmelanotide works as a substitute for MSH (Fig. 5.7). The first clinical setmelanotide study in patients with genetic obesity [34] described two adult patients with autosomal-recessive mutations in the POMC gene. The patients had a baseline BMI of 49.8 and 54.1 kg/m<sup>2</sup>, respectively. After 42 weeks, patient 1 had a BMI of 33.4 kg/m<sup>2</sup> and lost 51 kg (32.9% of her initial body weight). The second patient lost 20.5 kg (13.4% of her body weight) at 12 weeks of treatment. The hyperphagia was reduced in both patients during the treatment. There were no serious adverse events during the trial, but



interestingly the patients did notice that their hair and skin color darkened during the treatment. This might be due to stimulation of the melanocortin-1 receptor.

### 5.5.6 Obesity does not run in your family, the problem is that nobody runs in your family—obesity stigma and genetic counseling

Last but not the least, genetic counseling is a part of treatment for genetic obesity that should not be overlooked. Establishing a genetic diagnosis can end the “diagnostic odyssey,” the journey that patients and their families have to undertake to reach an etiological diagnosis. After the diagnosis, personalized treatment or specific guidance or support (e.g., in controlling hyperphagia) may become available. Besides that, a genetic diagnosis can reduce the obesity stigma, giving insight that obesity is not only a matter of poor lifestyle choices and little willpower. Establishing a genetic diagnosis may also support reproductive decision-making and help early intervention in other family members.

## 5.6 Summary

Obesity is the accumulation of body fat caused by an imbalance of energy intake and energy output. This balance is controlled by a complex neural system of which the most important processes take place in the hypothalamus. For a small percentage of people with obesity a genetic defect is the main cause of their obesity. Genetic obesity is often divided into syndromic obesity and nonsyndromic monogenic obesity. Syndromic obesity describes the patient that is not only obese but also has intellectual disability, congenital (organ) malformations, and/or dysmorphic features. The patients with nonsyndromic obesity mostly have one main manifestation of their genetic disease: obesity. The most important group of genetic obesity syndromes without intellectual disability is the syndromes caused by mutations in the leptin–melanocortin pathway. In these patients the hypothalamic satiety system is malfunctioning, resulting in early-onset obesity due to hyperphagia. MC4R deficiency is the most frequently identified cause of monogenic nonsyndromic obesity. The most common obesity syndrome with intellectual disability is PWS. In contrast to the leptin–melanocortin pathway–related obesity, PWS patients become obese later in childhood (> 5 years old). There is currently only one successful therapeutic agent on the market that targets a specific disruption in the leptin–melanocortin pathway: the administration of recombinant human leptin (metreleptin) to patients with leptin deficiency. Other drugs are underway, but a better understanding of the molecular and psychological mechanism behind food intake and the development of obesity is crucial before they will become commercially available. This could lead to understanding the large differences in the susceptibility for obesity between individuals, even in the same obesogenic environment. Hopefully more personalized or mechanism-based treatment options will evolve from further studying the extremely interesting field of obesity genetics.

## References

- [1] World Health Organization. Obesity and overweight. Retrieved from <<http://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight>>; 2017.
- [2] NCD Risk Factor Collaboration (NCD-RisC). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet* 2016;387(10026):1377–96. Available from: [https://doi.org/10.1016/S0140-6736\(16\)30054-X](https://doi.org/10.1016/S0140-6736(16)30054-X).
- [3] Calle EE, Thun MJ, Petrelli JM, Rodriguez C, Heath Jr. CW. Body-mass index and mortality in a prospective cohort of U.S. adults. *N Engl J Med* 1999;341(15):1097–105. Available from: <https://doi.org/10.1056/NEJM199910073411501>.
- [4] Reilly JJ, Methven E, McDowell ZC, Hacking B, Alexander D, Stewart L, et al. Health consequences of obesity. *Arch Dis Child* 2003;88(9):748–52.
- [5] Polderman TJ, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* 2015;47(7):702–9. Available from: <https://doi.org/10.1038/ng.3285>.
- [6] Stunkard AJ, Harris JR, Pedersen NL, McClearn GE. The body-mass index of twins who have been reared apart. *N Engl J Med* 1990;322(21):1483–7. Available from: <https://doi.org/10.1056/NEJM199005243222102>.
- [7] Farooqi IS, Keogh JM, Yeo GS, Lank EJ, Cheetham T, O'Rahilly S. Clinical spectrum of obesity and mutations in the melanocortin 4 receptor gene. *N Engl J Med* 2003;348(12):1085–95. Available from: <https://doi.org/10.1056/NEJMoa022050>.
- [8] Martinelli CE, Keogh JM, Greenfield JR, Henning E, van der Klaauw AA, Blackwood A, et al. Obesity due to melanocortin 4 receptor (MC4R) deficiency is associated with increased linear growth and final height, fasting hyperinsulinemia, and incompletely suppressed growth hormone secretion. *J Clin Endocrinol Metab* 2011;96(1):E181–8. Available from: <https://doi.org/10.1210/jc.2010-1369>.



- [9] Walters RG, Jacquemont S, Valsesia A, de Smith AJ, Martinet D, Andersson J, et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 2010;463(7281):671–5. Available from: <https://doi.org/10.1038/nature08727>.
- [10] Doche ME, Bochukova EG, Su HW, Pearce LR, Keogh JM, Henning E, et al. Human SH2B1 mutations are associated with maladaptive behaviors and obesity. *J Clin Invest* 2012;122(12):4732–6. Available from: <https://doi.org/10.1172/JCI62696>.
- [11] Ren D, Zhou Y, Morris D, Li M, Li Z, Rui L. Neuronal SH2B1 is essential for controlling energy and glucose homeostasis. *J Clin Invest* 2007;117(2):397–406. Available from: <https://doi.org/10.1172/JCI29417>.
- [12] Keith SW, Redden DT, Katzmarzyk PT, Boggiano MM, Hanlon EC, Benca RM, et al. Putative contributors to the secular increase in obesity: exploring the roads less traveled. *Int J Obes (Lond)* 2006;30(11):1585–94. Available from: <https://doi.org/10.1038/sj.ijo.0803326>.
- [13] van Rossum EF. Obesity and cortisol: new perspectives on an old theme. *Obesity (Silver Spring)* 2017;25(3):500–1. Available from: <https://doi.org/10.1002/oby.21774>.
- [14] Montague CT, Farooqi IS, Whitehead JP, Soos MA, Rau H, Wareham NJ, et al. Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* 1997;387(6636):903–8. Available from: <https://doi.org/10.1038/43185>.
- [15] Clement K, Vaisse C, Lahlou N, Cabrol S, Pelloux V, Cassuto D, et al. A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature* 1998;392(6674):398–401. Available from: <https://doi.org/10.1038/32911>.
- [16] Yeo GS, Farooqi IS, Aminian S, Halsall DJ, Stanhope RG, O'Rahilly S. A frameshift mutation in MC4R associated with dominantly inherited human obesity. *Nat Genet* 1998;20(2):111–12. Available from: <https://doi.org/10.1038/2404>.
- [17] Vaisse C, Clement K, Guy-Grand B, Froguel P. A frameshift mutation in human MC4R is associated with a dominant form of obesity. *Nat Genet* 1998;20(2):113–14. Available from: <https://doi.org/10.1038/2407>.
- [18] Krude H, Biebermann H, Luck W, Horn R, Brabant G, Gruters A. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat Genet* 1998;19(2):155–7. Available from: <https://doi.org/10.1038/509>.
- [19] Miller JL, Lynn CH, Driscoll DC, Goldstone AP, Gold JA, Kimonis V, et al. Nutritional phases in Prader-Willi syndrome. *Am J Med Genet A* 2011;155A(5):1040–9. Available from: <https://doi.org/10.1002/ajmg.a.33951>.
- [20] Poxe-Wolf J, Lam BY, Larder R, Tadross J, Rimmington D, Bosch F, et al. Hypothalamic loss of Snord116 recapitulates the hyperphagia of Prader-Willi syndrome. *J Clin Invest* 2018;128(3):960–9. Available from: <https://doi.org/10.1172/JCI97007>.
- [21] Forsythe E, Beales PL. Bardet-Biedl syndrome. *Eur J Hum Genet* 2013;21(1):8–13. Available from: <https://doi.org/10.1038/ejhg.2012.115>.
- [22] Sheffield VC. The blind leading the obese: the molecular pathophysiology of a human obesity syndrome. *Trans Am Clin Climatol Assoc* 2010;121:172–81 discussion 181–172.
- [23] Grace C, Beales P, Summerbell C, Jebb SA, Wright A, Parker D, et al. Energy metabolism in Bardet-Biedl syndrome. *Int J Obes Relat Metab Disord* 2003;27(11):1319–24. Available from: <https://doi.org/10.1038/sj.ijo.0802420>.
- [24] Wabitsch M, Funcke JB, Lennerz B, Kuhnle-Krahl U, Lahr G, Debatin KM, et al. Biologically inactive leptin and early-onset extreme obesity. *N Engl J Med* 2015;372(1):48–54. Available from: <https://doi.org/10.1056/NEJMoa1406653>.
- [25] Styne DM, Arslanian SA, Connor EL, Farooqi IS, Murad MH, Silverstein JH, et al. Pediatric obesity-assessment, treatment, and prevention: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab* 2017;102(3):709–57. Available from: <https://doi.org/10.1210/jc.2016-2573>.
- [26] Buchwald H, Avidor Y, Braunwald E, Jensen MD, Pories W, Fahrenbach K, et al. Bariatric surgery: a systematic review and meta-analysis. *J Am Med Assoc* 2004;292(14):1724–37. Available from: <https://doi.org/10.1001/jama.292.14.1724>.
- [27] Eldar S, Heneghan HM, Brethauer SA, Schauer PR. Bariatric surgery for treatment of obesity. *Int J Obes (Lond)* 2011;35(Suppl. 3):S16–21. Available from: <https://doi.org/10.1038/ijo.2011.142>.
- [28] Censani M, Conroy R, Deng L, Oberfield SE, McMahon DJ, Zitsman JL, et al. Weight loss after bariatric surgery in morbidly obese adolescents with MC4R mutations. *Obesity (Silver Spring)* 2014;22(1):225–31. Available from: <https://doi.org/10.1002/oby.20511>.
- [29] Bonnefond A, Keller R, Meyre D, Stutzmann F, Thuillier D, Stefanov DG, et al. Eating behavior, low-frequency functional mutations in the melanocortin-4 receptor (MC4R) gene, and outcomes of bariatric operations: a 6-year prospective study. *Diabetes Care* 2016;39(8):1384–92. Available from: <https://doi.org/10.2337/dc16-0115>.
- [30] Le Beyec J, Cugnet-Anceau C, Pepin D, Alili R, Cotillard A, Lacorte JM, et al. Homozygous leptin receptor mutation due to uniparental disomy of chromosome 1: response to bariatric surgery. *J Clin Endocrinol Metab* 2013;98(2):E397–402. Available from: <https://doi.org/10.1210/jc.2012-2779>.
- [31] Alqahtani AR, Elahmedi MO, Al Qahtani AR, Lee J, Butler MG. Laparoscopic sleeve gastrectomy in children and adolescents with Prader-Willi syndrome: a matched-control study. *Surg Obes Relat Dis* 2016;12(1):100–10. Available from: <https://doi.org/10.1016/j.soard.2015.07.014>.
- [32] Farooqi IS, Jebb SA, Langmack G, Lawrence E, Cheetham CH, Prentice AM, et al. Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N Engl J Med* 1999;341(12):879–84. Available from: <https://doi.org/10.1056/NEJM199909163411204>.
- [33] Chen KY, Muniyappa R, Abel BS, Mullins KP, Staker P, Brychta RJ, et al. RM-493, a melanocortin-4 receptor (MC4R) agonist, increases resting energy expenditure in obese individuals. *J Clin Endocrinol Metab* 2015;100(4):1639–45. Available from: <https://doi.org/10.1210/jc.2014-4024>.
- [34] Kuhnen P, Clement K, Wiegand S, Blankenstein O, Gottesdiener K, Martini LL, et al. Proopiomelanocortin deficiency treated with a melanocortin-4 receptor agonist. *N Engl J Med* 2016;375(3):240–6. Available from: <https://doi.org/10.1056/NEJMoa1512693>.

## Guide to further reading: articles

For background information on case 2 Kleinendorst L, van Haelst MM, van den Akker ELT. Young girl with severe early-onset obesity and hyperphagia. *BMJ Case Rep* 2017. doi:10.1136/bcr-2017-221067.

For more information on the sensory input to the brain in relation to food intake and appetite Berthoud HR, Morrison C. The brain, appetite, and obesity. *Annu Rev Psychol* 2008;59:55–92. doi:10.1146/annurev.psych.59.103006.093551.

For information on stress and obesity van der Valk ES, Savas M, van Rossum, EFC. Stress and obesity: are there more susceptible individuals?. *Curr Obes Rep* 2018;7(2):193–203. doi:10.1007/s13679-018-0306-y.

- For information on Prader-Willi syndrome* Cassidy SB, Schwartz S, Miller JL, Driscoll DJ. Prader-Willi syndrome. *Genet Med* 2012;14(1):10–26. doi:10.1038/gim.0b013e31822bead0.
- For information on ciliopathies and obesity* de Vries TI, van Haelst MM. Ciliary disturbances in syndromal and non-syndromal obesity. *J Pediatr Genet* 2014;3(2):79–88. doi:10.3233/PGE-14085.
- For information on the mirror phenotype of 16p11.2 deletions and duplications* Jacquemont S, Reymond A, Zufferey F, Harewood L, Walters RG, Kutalik Z, et al. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* 2011;478(7367):97–102. doi:10.1038/nature10406.

### **Online material**

- For factsheets, publications and statistics on obesity* <[www.who.int/topics/obesity/en/](http://www.who.int/topics/obesity/en/)>.
- The World Obesity Image Bank* <[www.imagebank.worldobesity.org](http://www.imagebank.worldobesity.org)>.
- For information on 16p11.2 microdeletion syndrome* <[www.rarechromo.org/disorder-guides](http://www.rarechromo.org/disorder-guides)>.

# Molecular dysmorphology

Koumudi Godbole

Clinical Geneticist, Department of Genetics, Deenanath Mangeshkar Hospital and Research Center, Pune, India

## 6.1 Introduction

Dysmorphology [“dys” (disordered or abnormal) and “morph” (shape, abnormal) or disordered shape] is a favorite word of clinical geneticists, first coined by Smith [1]. It implies a study of abnormal morphogenesis/development including physical defects of body and its structure, originating prior to birth. The term “dysmorphic” is used to describe individuals whose physical features such as face deviate from what is typically seen in the context of age, sex, and ethnic background.

Dysmorphism of an organ or body part may arise secondary to the following:

*Malformation*: structural/morphological defects of an organ or a larger body region resulting from an intrinsically abnormal developmental process.

*Deformation*: a distortion resulting from mechanical forces.

*Disruption*: destruction of an intrinsically normal tissue leading to an abnormal form, shape, or position of a part of the body.

A *sequence* is a pattern of multiple defects resulting from a single primary malformation [e.g., congenital talipes equino varus (CTEV)] and hydrocephalus resulting secondary to lumbar neural tube defect (NTD) while an *association* is a group of malformations occurring together more often than just by chance but without a yet known underlying etiology (e.g., VATER/VACTREL with vertebral, anal, tracheaesophageal, renal, and limb abnormalities). A *syndrome* is a pattern of features often with a unifying underlying cause (e.g., down syndrome due to trisomy of chromosome 21 or 21q containing the down syndrome critical region).

However, at times, it is difficult to determine whether a given anomaly is an intrinsic malformation or an extrinsic deformation/disruption. Also, with improved understanding of molecular basis, an association may get converted into a syndrome: the classical CHARGE association (coloboma of eye, choanal atresia, heart defects, retardation of growth and development, ear defects) is now known to be caused by mutations in CHD7 and SEMA3E genes converting it to CHARGE syndrome.

Before the advent of genomic tools the key diagnostic handle for genetic disorders was an in-depth clinical examination for dysmorphism and/or malformations and identifying unique characteristic features of a particular syndrome to offer relevant counseling and management, also known as “clinical” dysmorphology. Dysmorphism of external or internal body parts is often the first feature pointing to a genetic etiology as it is common to many clinically defined genetic syndromes. Even today, after over half a century from discovery of DNA and its structure, dysmorphism is one of the leading reasons to begin genetic/genomic investigations. However, it is important not to equate dysmorphism with a genetic etiology by default which could lead to a misdiagnosis.

### 6.1.1 Limitations of “clinical” dysmorphology and the newer dysmorphology tools

Diagnosing a rare dysmorphology syndrome is an integral part of a typical clinical genetics practice. Pattern recognition is a key skill in clinical dysmorphology aided by finding a match to the findings under consideration

in the published literature. However, absence of a previously described alike may not necessarily indicate the novelty of the syndrome because the matching process remains largely subjective [2]. Dysmorphology alone might be misleading as it is also subjective to experience of the geneticist, variability of phenotype, and phenocopies [3]. Academic forums examining such cases included dysmorphology meetings; peer-reviewed literature seeks to establish whether a particular syndrome is truly novel. This process lacks throughput, and it may take a long time before a novel recognizable syndrome is designated. Tools such as dysmorphology databases and computerized dysmorphology analysis may aid clinicians in reaching a diagnosis.

*Dysmorphology databases* are useful tools that require accurate identification of clinical signs as a handle to reach a diagnosis. These are typically equipped with patient photographs, characteristic radiological images, and allow filtering for essential and additional features, inheritance patterns, which might help in narrowing down on a smaller number of differentials. Since features are prelisted, a clinician can be reminded of other clinical features that are not so well-known and are missed. However, just keying some features in databases may not always lead to a diagnosis as there could be hundreds of syndromes with those features. It is not easy to identify an accurate diagnosis from background noise of other syndromes bearing the same clinical features and knowledge of genetic syndromes, which is often necessary to sieve out the conditions of interest.

*Computerized dysmorphology analysis:* There is an increasing recognition of use of 3D facial image analysis in identifying facial dysmorphology and related genetic disorders. Recent advances in computational models of facial dysmorphology can assist early recognition of characteristic face shape presented as frontal and supine photographs. Patients and control subjects are matched for gender and age. Software with specific mathematical programs have been developed, which extract and analyze information from certain landmarks (a grid of nodes placed on relevant facial structures) on these photographs and categorize them using a classification algorithm. An unknown subject is classified by assessing his/her similarity to case/control group within the database [4]. Although such efforts contribute to making the hazy field of subjective dysmorphology assessment more objective and user friendly, there is a lot of scope for improvement in computational dysmorphology. Faces might appear different in different ethnicities depending on their “normal” pattern, and they keep changing over time in both normal and syndromic individuals. Studies using subjective and objective statistical approaches have revealed contrasting age-related changes that either render the phenotype subtler or make it more apparent with age, making a clinical diagnosis challenging in older and younger individuals, respectively [5].

It is speculated that such advanced facial analysis might in future inform the strategy for molecular analysis and confirmation of a clinical diagnosis. Studies combining morphological and molecular analyses of atypical patients with a genetic condition can help identify the role of individual genes in different facets of the associated clinical pathologies.

### 6.1.2 Molecular dysmorphology

When the clinical practice of dysmorphology diagnosis merges with the research disciplines of developmental biology and genomics, the combination may be called “molecular dysmorphology.” This may lead to a unified diagnostic system by an optimal matching of the phenotype with molecular data providing advantages of both the disciplines [6].

The primary goal of this scientific discipline is delineation of the molecular basis of normal and abnormal development and in turn improvement of patient care through better understanding of fundamental pathways leading to abnormal development. Molecular dysmorphology relates to basic embryology that provides increasing correlation of specific embryological events with gene expression/function and fundamental cellular pathways that build organs and animals. Its connection to human disorders is possible because of remarkable functional conservation of many key functioning molecules and mechanisms between model systems and human development.

The journey of a “novel” malformation–dysmorphism syndrome often begins as a single case report with a “unique pattern” to a recurrent pattern “syndrome” that has an established genetic basis and a laboratory test, which may help in diagnosing individuals with atypical clinical features. ATR-X-alpha thalassemia is a typical example of this transition from diagnosis of thalassemia in some patients with mental retardation and specific dysmorphism [3]. With advent of molecular techniques, a subgroup of these patients was diagnosed with 16p deletion, and the remaining patients were eventually diagnosed to have alterations in the ATR-X gene located at Xq21.1. Thus the unusual finding of thalassemia with mental retardation has moved from a clinical observation made 25 years ago to a clearly defined, clinically identifiable disorder for which genetic testing is

now available. Many disorders, including Smith–Fineman–Myers syndrome, Xlmr–Hypotonic Facies syndrome, Carpenter–Waziri syndrome, Chudley–Lowry syndrome, Juberg–Marsidi syndrome, and Holmes–Gang syndrome with distinct clinical features but with consistent features of mental retardation–short stature are now known, all allelic to ATRX gene.

### 6.1.3 Molecular diagnosis and molecular medicine

Conventional Mendelian disorders result from point mutations in particular genes resulting from DNA replication or repair errors, while genomic disorders are thought to be a consequence of abnormal dosage of gene(s) located within a rearranged segment of the genome. These could be larger alterations in the genome: del/dup, inversions, and translocations that occur via recombination mechanisms. Molecular diagnosis of genomic rearrangements has seen a shift from cytogenetic techniques, such as G-banding to locus specific fluorescent in situ hybridization (FISH), chromosome painting, telomere FISH to array-CGH using BAC/PAC clones. The latter is a high throughput technology with potential to identify new genomic disorders or detecting submicroscopic rearrangements not visible by routine karyotyping. This may also detect reciprocal duplications of microdeletions that are presumably underascertained due to the mild degree or lack of appreciable phenotype. Next generation sequencing (NGS) has enabled probing into the whole exome (the coding exons together) and the whole genome (exons and introns) for sequence and/or copy number variation.

These powerful techniques can sometimes bypass the traditional sequence of clinical diagnosis first and then the specific genetic testing. Genome-wide copy number analysis and NGS have power to identify likely causal mutation (genotype) regardless the knowledge of the disease (phenotype). Matching a given phenotype with the novel candidate gene has been reported in certain genetic conditions, such as retinal dystrophy and neurocognitive disorders, especially in multiplex consanguineous families supporting a “genotype-first” approach. Such studies can sometimes provide a dual diagnosis and presence of more than one underlying causal mutation in consanguineous pedigrees. However the diagnostic yield of these powerful techniques remains lower than expected mainly because of the large number of rare variants and lack of functional studies. Research is being directed towards identifying tools that will help prioritize the most likely disease-causing mutations based on computational comparison of phenotypical abnormalities and novel approaches using gene networks and RNS-seq data. Exomizer, Phenoxome, Gene Network Assisted Diagnostic Optimization (GADO) are some examples of such efforts [7,8]. In the meanwhile it might be useful if studies with tentative genotype-phenotype links get published which would help in generation of such tools. This might lead to a trend of accelerating establishment of novel syndromes and their underlying genes and assist in upgrading “variants of uncertain significance” to pathogenic mutations.

Most patients in genetic clinics now get an objective etiological diagnostic test to support, confirm, or discover their diagnoses as various methods, including karyotyping, FISH, chromosomal and oligo-arrays, whole exome (and whole genome) sequencing, have now reached the doorsteps of clinicians. It is the clinician’s responsibility to keep up with these advances in order to choose an appropriate test that would be cost-effective and offer a specific molecular diagnosis to a patient with suspected dysmorphism–malformation syndrome.

The making of an accurate diagnosis, thereby bringing closure and offering a possible definitive treatment and/or genetic counseling, is a typical sequence of events when dealing with patients of intractable disorders without a classical unique phenotype. Joshi et al. [9] have reported a diagnostic odyssey of a nonspecific phenotype of early onset epileptic encephalopathy where average age of starting investigations was six months and average age at diagnosis was 95 months to about 8 years of diagnostic process (commonly including EEG, MRI of brain, tests for IEM, muscle/nerve/skin biopsy, EMG, CSF studies, and genetic studies including specific single gene sequencing such as Rett/channelopathy genes). Molecular diagnostic yield in these patients using genomic tests was over 40%, which helped in providing families with better understanding of the disorder and appropriate genetic counseling.

The knowledge of underlying molecular defects also improves the understanding of possible molecular targets and complemented with genetic engineering may provide a molecular solution to ameliorate the genetic condition.

## 6.2 Clinical cases and molecular basis

The initial decades of clinical genetics were dominated by dysmorphologists, and genetic conditions were described based on the phenotype. With increasing understanding of molecular biology and rapidly advancing



molecular techniques, the field of clinical genetics and dysmorphology has been undergoing transition into molecular dysmorphology. Molecular discoveries are providing newer definitions to many previously clinically delineated syndromes often by lumping or splitting. We now know that certain genetic conditions with a specific phenotype are in fact genetically heterogeneous, while certain phenotypically distinct syndromes may result from different pathogenic alterations in the same gene. We will discuss a few genetic conditions that demonstrate different molecular aspects of dysmorphology in the following section.

## 6.2.1 Genetic heterogeneity

### 6.2.1.1 The splitting

Here we will discuss some examples of a seemingly one specific phenotype with genetic heterogeneity.

#### Holoprosencephaly

Holoprosencephaly (OMIM 236100) is the commonest developmental defect of forebrain and is characterized by inadequate or absent midline division of the developing forebrain into cerebral hemispheres. Its incidence is reportedly 1:250 in conceptuses and about 1:10,000 at term, and it is associated with about 80% of midline facial defects.

Holoprosencephaly represents a continuum of hemispheric separation abnormalities and is divided into four main subtypes: lobar, semilobar, lobar holoprosencephaly, and middle hemispheric variant with decreasing severity. Associated facial defects vary in severity: from hypotelorism to cyclopia, frontonasal anomaly from reduced distance between the nares, single nostril to proboscis, a nose-like elongated structure, midline cleft lip with or without palate, midface hypoplasia, microretrognathia, and a single maxillary central incisor. Microforms, including hypotelorism, midface hypoplasia, and single maxillary central incisor, are milder and nonlethal, while severe forms, such as cyclopia and proboscis, are generally lethal.

Holoprosencephaly is caused by both genetic and/or environmental factors and includes a spectrum of defects ranging from mild to severe. It is a classic example of a complex genetic trait with “pseudo”-autosomal dominant transmission showing incomplete penetrance and variable expressivity. Although there is some correlation between severity of midline facial defects and severity of brain defects, neither the associated brain abnormality nor the underlying genetic/environmental causative factor can be accurately predicted by overtly visible facial defect. The same causative factor can potentially present with any of the holoprosencephaly spectrum anomalies, and several molecular models have been proposed to explain unclear genotype–phenotype correlation. Variable gene dosage, gene–gene interaction, which may include two genes known to cause holoprosencephaly or other genes involved in forebrain developmental pathways, genetic modifiers, gene–environment interaction, and stochastic/epigenetic contribution, have been suggested and mostly support a “multiple hit” hypothesis. [Table 6.1](#) depicts the common causes of holoprosencephaly in humans [10].

Environmental factors, including fetal exposure to maternal diabetes, retinoic acid, alcohol, are also associated with holoprosencephaly. Retinoic acid plays a critical role in anteroposterior patterning of the central nervous system, and there is a tight regulation of retinaldehyde dehydrogenase conversion of vitamin A to all-trans retinoic acid as well as CYP26A1, a retinoic acid degrading enzyme. TGIF regulates the expression of genes, controlling both its synthesis and degradation. Mechanisms of alcohol toxicity are not well understood although several potential mechanisms of ethanol teratogenesis require its oxidative metabolism. EtOH and its acetaldehyde might act as competitive inhibitors of ADH/ALDH enzymes involved in retinoic acid production, which is required for normal developmental patterning. Reactive oxygen species produced during EtOH metabolism and direct effects of alcohol in perturbing cellular membranes and inhibiting functions of certain membrane proteins, such as cell adhesion molecules, might additionally contribute to ethanol toxicity [11].

Maternal diabetes and consumption of cholesterol lowering drugs have also been associated with holoprosencephaly. Infants of diabetic mothers have almost 1% (200 folds) increased risk of having holoprosencephaly, while statins probably lead to altered biogenesis of sonic hedgehog (SHH) and signal transduction in SHH-mediated cells due to reduced cholesterol levels.

[Table 6.1](#) depicts common genetic etiology of human holoprosencephaly.

Apart from the above genes mainly contributing to holoprosencephaly, pathogenic variants in other genes, namely, CDON, DISP1, DLL1, FGF8, GAS1, FOXH1, NODAL, and TDGF1, involved directly or indirectly through above major contributor genes, have been reported in persons with holoprosencephaly spectrum anomalies.



**TABLE 6.1** Common genetic causes of human holoprosencephaly.

Etiology	Genetic condition	Inheritance pattern	Recurrence risk
<b>Chromosomal</b>			
Chromosomal	Trisomy 13	De novo or inherited from a translocation carrier parent	Low or up to 50% in the case of parental translocation
	Trisomy 18	De novo or inherited from a translocation carrier parent	Low or up to 50% in the case of parental translocation
	Triploidy	De novo	Low
<b>Single gene defect</b>			
<b>Syndromic</b>			
Single gene disorder	Smith–Lemli–Opitz syndrome/7DHC	Autosomal recessive	25%
	Pallister–Hall syndrome/GLI3	De novo/sporadic	Low
	Rubinstein–Taybi syndrome/CREBP	Often de novo	Low
<b>Nonsyndromic</b>			
	Additional reported features		
SHH	Hydronephrosis	Often de novo/autosomal dominant	Low or up to 50% if inherited from a parent-incomplete penetrance
ZIC2	Bitemporal narrowing, neural tube defects		
SIX3	Endocrine dysgenesis		
GLI2	Postaxial polydactyly		
PTCH1	Macrocephaly, panhypopituitarism, omphalocele		

SHH, Sonic hedgehog.

It is important to understand the complex etiology of holoprosencephaly in order to initiate investigations in the proband and offer genetic counseling regarding reproductive risks for the family members. Typically, a high-resolution karyotype or array CHG as the first test to detect cytogenetic abnormalities, followed by the second tier tests, such as sequencing of the most common holoprosencephaly genes (SHH, ZIC2, SIX3) or specific gene for the suspected genetic syndrome, are generally recommended. NGS for holoprosencephaly-related genetic panels may follow if necessary [12].

Genetic counseling depends on the chromosomal or genetic etiology involved. As the presence of a mutation in a particular gene does not predict the severity even within the same family, genetic counseling and use of genetic tests for prenatal diagnosis should be used with caution. Fetal ultrasound and brain MRI are useful tools in the prenatal diagnosis of holoprosencephaly when molecular diagnosis is not known.

### Split hand–foot malformation

Split hand–foot malformation (SHFM) is a congenital limb defect mainly affecting central rays of the hands and/or feet and accounts for about 10%–15% of all limb malformations. It may be an isolated finding or may occur as a part of a syndrome. SHFM has a spectrum of severity from mild defects, including hypoplasia of a single phalanx or syndactyly, to severe defects with aplasia of one or more central digits: lobster claw anomaly or monodactyly, where both central and preaxial rays are affected.

Like holoprosencephaly, there is a significant inter and intrafamilial variable expressivity within families [13]. Although most cases are isolated, sporadic and show autosomal dominant mode of inheritance, autosomal and X-linked recessive inheritance and syndromic occurrence of SHFM are also known. In addition, sex bias with increased transmission of genetic alteration from affected father to son has been reported for dup17p13.3-associated

**TABLE 6.2** Common causes of split hand–foot malformation.

Type	Genetic abnormality	Phenotype	Inheritance
SHFM1	7q21.3-q22.1 rearrangement	EEC, hearing loss, mental retardation	AD
	DLX5 mutations		AR
SHFM2	Xq26	SHFM, syndactyly hypoplastic metacarpals and phalanges	XR
SHFM3	Dup10q24	SHFM	AD
SHFM4	TP63	SHFM or syndromic: EEC, ADULT, LADD, CHARGE, VATER	AD
SHFM5	Del 2q31	SHFM	AD
SHFM6	WNT10B	SHFM, tibial aplasia/hypoplasia	AR
SHFLD1	1q42.2-q43	Hypoplasia/aplasia of forearm and leg bones	AD
SHFLD2	6q14.1	Long-bone deficiency	
SHFLD3	Dup 17p13.3/BHLHA9	SHFM, tibial aplasia /hypoplasia	AD

AD, Autosomal dominant; AR, autosomal recessive; SHFL, split hand–foot malformation with long-bone deficiency; SHFM, split hand–foot malformation; XR, X-linked recessive.

SHFM. Females are usually asymptomatic carriers, but if affected, they have the most severe phenotype [14]. Table 6.2 lists the currently known causes of SHFM that account for about 50% of SHFMs.

Appropriate genetic counseling can only be provided based on the underlying genetic cause. Of the known causes, array CGH may detect about one-third of cases with chromosomal del/dups, duplication of 10q24 being the commonest. This may be followed by TP63 sequencing or by a SHFM gene panel by NGS.

### 6.2.1.2 Syndromes with genetic connections: the lumping

There have always been well-defined classical genetic syndromes with some overlapping clinical features. These have eventually been grouped together after understanding common gene alteration or alteration in genes involved in a particular pathway. New terminologies have now been coined to address the group of disorders due to a particular gene/biological system–pathway defect, for example, RASopathies or Transcriptomopathies. Such examples will be discussed in this section to demonstrate a common molecular genetic theme underlying apparently distinct clinical phenotypes/syndromes. It is speculated that deciphering pathway interconnection of clinically similar syndromes may provide common targets useful for developing new therapeutics.

#### RASopathies

RASopathies comprise a group of disorders clinically characterized by short stature, heart defects, facial dysmorphism, varying degrees of intellectual disability, and cancer predisposition. They are caused by germline variants in genes encoding key components or modulators of the highly conserved RAS–MAPK (mitogen-activated protein kinase) signaling pathway that leads to dysregulation of cell signal transmission. RASopathies include neurofibromatosis type 1, Noonan syndrome, Leopard syndrome, hereditary gingival fibromatosis type 1, capillary malformation-AV malformation syndrome, cardiofaciocutaneous syndrome, and autoimmune lymphoproliferative syndrome [15,16]. Although each one of them has certain unique phenotypic features, they share certain clinical signs, such as craniofacial anomalies, heart defects, short stature, neurocognitive deficits, cutaneous and musculoskeletal abnormalities, and a predisposition to malignancies, in common. An overlap of many of these features has been described previously, which makes a specific clinical diagnosis and genotype–phenotype correlation rather challenging in RASopathies. Table 6.3 describes various syndromes grouped together as RASopathies [17].

#### Laminopathies

Laminopathies are a group of rare genetic disorders caused by mutations in genes encoding proteins of the nuclear lamina. Diseases caused by mutations in LMNA include autosomal dominant Emery–Dreifuss muscular dystrophy and related myopathies, Dunnigan-type familial partial lipodystrophy, Charcot–Marie–Tooth disease type 2B1 and developmental and accelerated aging disorders. These mutations cause four different phenotypes

**TABLE 6.3** The RASopathies.

Syndrome	Gene(s)	Typical features
Neurofibromatosis	NF1	Neurofibromas, Lisch nodules, cutaneous lesions such as café au lait spots, predisposition to malignancies
Noonan syndrome	PTPN11, KRAS, SOS1, RAF1, NRAS, BRAF, SHOC2, CBL, RIT1	Postnatal growth retardation, distinct facial dysmorphism, cardiac defects, cognitive impairment, predisposition to malignancies
Leopard syndrome	PTPN11, BRAF, RAF1	Lentigenes, café au lait spots, HOCM, electrocardiographic conduction anomalies, pulmonary stenosis, abnormal genitals, growth retardation, and sensorineural deafness
Neurofibromatosis1-like syndrome/Legius syndrome	SPRED1	Mild NF1 phenotype with pigmentary changes, macrocephaly, learning difficulties, and lipomatosis
RAS-associated autoimmune lymphoproliferative disorder (ALPS4)	NRAS, KRAS	Autoimmune pancytopenia, defective lymphocyte proliferation, hypergammaglobulinemia, chronic lymphadenopathy, predisposition to malignancies
Noonan syndrome-like disorder with loose anagen hair	SHOC2, PPP1CB	Thin sparse slow growing hair, macrocephaly, developmental delay
Costello syndrome	HRAS	Fetal overgrowth, macrocephaly, short stature, coarse facial features, cutis laxa, slow growing, sparse hair with abnormal texture, papillomas, vascular lesions, cognitive deficits, predisposition to malignancies
Cardiofaciocutaneous syndrome	BRAF, MP2K1, MP2K2, KRAS	Short stature, macrocephaly, lentigenes, cardiac defects, facial dysmorphism, lax joints, sparse, slow growing hair

with variable overlap as seen in above syndromes, namely, striated muscle disease, lipodystrophy, peripheral neuropathy, and accelerated aging [18].

Duplication in LMNB1 gene causes autosomal dominant leukodystrophy, and mutations in LMNB2 are associated with acquired partial lipodystrophy. In addition, mutations in genes encoding lamin-associated integral inner nuclear membrane proteins cause X-linked Emery–Dreifuss muscular dystrophy, sclerosing bone dysplasias, HEM/Greenberg skeletal dysplasia and Pelger–Huet anomaly.

It is rather intriguing how different mutations in this single gene can cause different, often system-specific, disease phenotypes. Striated muscle appears to be sensitive to both mutations and reduced expression of A-type lamins resulting in structural disruption of nuclear lamina and nuclear envelope. This “weakened” lamina may lead to overall loss of a cell’s ability to withstand stress-induced damage, which may be of critical significance in skeletal and cardiac muscle contraction. Many LMNA mutations causing striated muscle disease or loss of A-type lamins from cells lead to a redistribution of emerin from the nuclear envelope to the bulk endoplasmic reticulum. The farnesyl group of unprocessed lamin A and truncated prelamin A is likely responsible for the perturbed DNA repair and genomic instability. Lamin A may negatively regulate adipocyte differentiation and accumulation of unprocessed prelamin A reportedly blocks adipocyte differentiation linking lipodystrophy phenotype to laminopathies.

A multidisciplinary medical team is required to manage laminopathies. Surgical and conservative procedures may prevent progression or improve contractures and scoliosis. Implantation of a cardioverter defibrillator has been recommended to prevent sudden death due to lethal tachyarrhythmias. New personalized therapeutic approaches, including exon skipping, RNAi, rapamycin—alone or in combination with all-trans retinoic acid and Temozolomide—a rapamycin analog are being investigated to treat laminopathies and will be discussed later in this chapter.

### Ciliopathies

Defects of the primary cilium, a microtubule-based projection of nondividing cells, are now known to cause clinically recognizable developmental syndromes together called “ciliopathies” whose spectrum involves nearly every body organ. Vesicle trafficking is the major process by which components are acquired for cilium formation and maintenance [19]. Mutations occurring in genes contributing to the overall vesicle trafficking to the primary cilium and/or affecting ciliary assembly might lead to more severe symptoms, while those involved in the

**TABLE 6.4** Ciliopathies revealing overlapping phenotype and genetic heterogeneity.

Ciliopathy	Phenotype	Gene(s)
BBS	Obesity, retinitis pigmentosa, polydactyly, genitourinary anomalies, mental retardation	BBS: 1/2/4/5/7/7/10/12, ARL6, MKKS, TCT8, TRIM32, WDPCP, LZTFL1, BBIP1
MGS	Polycystic kidneys/renal agenesis, polydactyly, occipital encephalocele, and brain, cardiac, urogenital defects	MKS1, MKS7, TMEM216/231, TMEM138, TMEM 67, TCTN2, CC2D2A, CEP290, RPGRIP1L, B9D1, B9D2
Joubert syndrome (cerebro–retino–renal syndrome)	Multisystem anomalies of eye, kidney, brain, heart, skeletal and mental disability, phenotype variation with underlying genetic etiology	INPP5E, AHI1, ARL13B, OFD1, KIF7, TCTN1, CEP41, C5ORF42, ZNF423, NPHP1, TCT21B, CEP164, CEP290, MRE11, TMEM216/231, TMEM138, TMEM 67, TCTN2, RPGRIP1L, B9D1, B9D2
Nephronophthisis	Nephronophthisis, end stage renal disease, growth retardation, anemia	INV, GLIS2, NEK8, IFT172, IFT144, ZNF423, NPHP1, TCT21B, CEP164, CEP290, MRE11, RPGRIP1L, TMEM 67, TCTN2, CC2D2A, CEP290, ANK6, NPHP18, NPHP4, IQCB1, SDCCAG8, MKS7
Senior Loken (renal–retinal) syndrome	Tapetoretinal degeneration, flat ERG, nephronophthisis, anemia	NPHP4, IQCB1, SDCCAG8, MKS7, CEP290

BBS, Bardet–Biedl syndrome; MGS, Meckel–Gruber syndrome.

transport of only certain substances might result in milder phenotypes. A broad rather than discrete expression of individual ciliopathy genes, stochastic events, and other genetic modifiers might contribute to this wide spectrum of clinical features. Shaheen et al. [20] reported a trend where genes that cluster in the same ciliary component caused a similar phenotype. Ciliopathies are typically autosomal recessive disorders arising from mutations in more than 50 genes, and their phenotype is a continuum spanning from embryonic lethality to isolated late onset retinal degeneration. Table 6.4 enumerates the known ciliopathies and genes involved.

### Filaminopathies

Filamins are actin-binding proteins that interact with multiple receptors and intracellular proteins. They stabilize the three-dimensional actin webs linking them to cell membranes and are important in cytoskeleton-dependent cell proliferation, differentiation, and migration [21]. Filamins A, B, and C (FLNA, FLNB, and FLNC) have been associated with human disorders mainly including skeleton and musculature-including cardiac muscle. Filamins A and C disorders are transmitted in recessive as well as dominant fashion, while the two known FLNC disorders show autosomal dominant inheritance pattern. While the null mutations in FLNA are commonly associated with disordered neuronal migration, cardiac function and connective tissue integrity, the missense mutations may lead to multisystemic affection, predominantly involving the skeleton. Three types of FLNC mutations have been reported: (1) mutations leading to the expression of misfolded FLNC, (2) mutations that do not affect protein solubility properties but give rise to a toxic gain of function by altering ligand-binding properties, and (3) mutations causing a premature stop codon and concomitant nonsense-mediated decay, resulting in haploinsufficiency. While the first type of mutations results in protein aggregation and subsequent impairment of protein homeostasis with a typical myofibrillar myopathy phenotype, the latter two types of mutations result in distal myopathy with no protein aggregates [22]. A wide phenotypic spectrum of filaminopathies is depicted in Table 6.5.

### 6.2.2 Epigenetic mechanisms and transcriptomopathies

Dysregulation of transcription during the complex orchestration during human embryogenesis is a known cause of many developmental syndromes. The transcription machinery including chromatin remodelers and histone-modifying enzymes is crucial in development and their disruption may lead to disorders of transcriptional regulation or transcriptomopathies [23]. Although there is a wide phenotypic spectrum, there are also some shared features besides cognitive impairment. Disruption of chromatin-associated transcription machinery accounts for the phenotypic overlap of Cornelia de Lange with KBG syndrome characterized by macrodontia of

**TABLE 6.5** The wide phenotypical spectrum of filaminopathies.

Gene (locus)	Disorder	Inheritance
FLNA (Xq28)	Frontometaphyseal dysplasia type 1	XLR
	Chronic idiopathic intestinal pseudoobstruction/congenital short bowel syndrome	XLR
	Cardiac valvular dysplasia	XLR
	Otopalatodigital syndrome, type 1 and 2	XLD
	Periventricular nodular heterotopias 1	XLD
	Melnick–Needles syndrome	XLD
	Terminal osseous dysplasia	XLD
	FG syndrome 2	XL
FLNB (3P14.3)	Spondylocarpotarsal synostosis syndrome	AR
	Atelosteogenesis I and III	AD
	Larsen syndrome	AD
	Boomerang dysplasia	AD
FLNC (7q32.1)	Familial restrictive cardiomyopathy type 5	AD
	Myofibrillar myopathy type 5	AD

AD, autosomal dominant; AR, Autosomal recessive; FLNA, Filamin A; FLNB, Filamin B; FLNC, Filamin C; XL, X linked; XLD, X linked dominant; XLR, X linked recessive.

the upper central incisors, distinctive craniofacial findings, short stature, skeletal anomalies, and neurologic involvement that includes global developmental delay, seizures, and intellectual disability [24].

Similarly, there is a significant overlap between CHARGE syndrome (OMIM 608892: Coloboma, heart defects, atresia of choanae, retardation of growth and development, genital hypoplasia and ear anomalies including deafness and vestibular disorders) and Kabuki syndrome (KS) (OMIM 147920: typical facial gestalt with growth retardation, heart defects, hearing loss, and intellectual disability). Clinical distinction, until the typical “kabuki” facies become apparent, could be difficult, and about 70% of individuals with typical/suspected CHARGE or KS are identified to have a pathogenic CHD7 or KMT2D mutation, respectively. These genes have a functional correlation via their interaction with the WAR complex (WDR5, RBBP5, and ASH2), and they might regulate a common set of genes. Their DNA methylation pattern appears to relate to genes involved in embryonic development of cell types. They have also been linked to expression of various homeobox-containing genes that encode highly conserved transcription factors expressed in a spatially and temporally regulated manner during development. DNA methylation appears to be a crucial element in developmental regulation of HOX activity postnatally. Some of the overlapping features of CHARGE and KSs, such as growth deficiency, skeletal and limb anomalies, renal dysgenesis, and neural development, might be mediated by reduced expression of HOXA5 [25]. Examples of malformation syndromes due to transcription dysregulation are listed in Table 6.6.

### 6.2.3 Spliceopathies

Disorders characterized by dysregulation of splicing regulating factors are collectively called “spliceopathies.” These disorders are suspected to result not from deficient protein encoded by the concerned gene but by toxic effects of the abnormally spliced RNA. While these mechanisms are only recently being discovered, their consequences are likely to produce the following detrimental effects: (1) RNA gain of toxic function, where the unstable and dynamic expansion of microsatellite sequence repeats in a noncoding region of mRNA, leads to intranuclear accumulation of mutated transcripts and the missplicing of numerous transcripts, (2) generation of a “sink” for RNA-binding proteins by increasing the mass of target RNA per nucleus and increasing avidity of RNA-protein interactions. Myotonic dystrophy type 1 (expanded CTG repeats in DMPK1 gene) and 2 (expanded CCTG repeats in ZNF9 gene), Fragile X-associated Tremor/Ataxia syndrome, spinocerebellar ataxia type 8 are considered to be some examples of spliceopathies. Spliceopathies are considered attractive targets for therapy as most of them do not directly cause absence or dysfunctional proteins [26].

**TABLE 6.6** Malformation syndromes with transcriptional dysregulation.

Transcriptional regulation and gene(s)	Syndrome	Clinical features
<i>Histone modification</i>		
CBP, EP300	Rubinstein–Taybi syndrome	Downslanting palpebral fissures, beaked nose, broad thumbs and big toes, short stature, and developmental delay
KMT2A	Wiedemann–Steiner syndrome	Hypertelorism, downslanting palpebral fissures, hypertrichosis, and developmental delay
KMT2D, KDM6A	Kabuki syndrome	Long palpebral fissures, eversion of lateral third of lower eyelids, skeletal anomalies, fetal fingertip pads, developmental delay, poor growth
KAT6B	Say–Barber–Biesecker syndrome/Young–Simpson syndrome	Blepharophimosis, developmental delay, hypotonia, joint stiffness, cryptorchidism, hypothyroidism, and patellar hypoplasia/agenesis
ANKRD11	KBG syndrome	Macrodontia of the upper central incisors, anteverted nares, short stature, skeletal anomalies, and developmental delay
<i>Chromatin remodeling</i>		
ARID1A, ARID1B, SMARCA4, SMARCB1, SMARCE1	Coffin–Siris syndrome	Coarse facies with bushy eyebrows, large mouth, aplasia/hypoplasia of the distal phalanx or nail, developmental delay, hypotonia
SMARCA2	Nicolaides–Baraitser syndrome	Coarse facial features, sparse scalp hair, interphalangeal joint swelling, microcephaly, seizures, and developmental delay
CHD7	CHARGE syndrome	Coloboma, heart defects, choanal atresia, retarded growth and development, genital abnormalities, and ear anomalies, facial asymmetry
<i>Mediator protein</i> MED12	FG syndrome	Developmental delay, macrocephaly with prominent forehead, hypotonia, corpus callosum abnormalities, and broad thumbs and big toes
<i>Cohesin complex-related</i> NIPBL, SMC1A, SMC3, HDAC8	Cornelia/Brachman de lange syndrome	Synophrys, arched eyebrows, long philtrum, developmental delay, growth retardation, hirsutism, and upper limb abnormalities
<i>Transcriptional elongation</i> AFF4	CHOPS syndrome	Developmental delay, heart defect, obesity, pulmonary abnormalities, short stature, and skeletal anomalies

## 6.3 Molecular diagnosis and therapy

Genetics is often considered as a peripheral than a core medical specialty, primarily due to the rarity of genetic disorders and limited therapeutic options. Traditionally genetic disorders are managed with supportive treatment using standard pharmacological agents. While bone marrow, organ, or hemopoietic stem cell transplant have been traditional ways of introducing cells with normal genetic makeup, in a hope to produce a normal protein product, gene therapy might be able to introduce a normal copy of the gene to restore the function of the protein.

Understanding the molecular basis of genetic disorders has brought the possibility of treating a genetic defect by a genetic medicine very close to reality. Molecular therapies aim at correcting the unique genetic abnormality in question and are a remarkable step toward personalized medicine. Many such molecular medicines are yet in the clinical trial phase and are surrounded by controversies about their long-term safety, efficacy, as well as ethical issues of creating a distinction between the rich and the poor due to extraordinary costs involved in conducting such trials and possible discrepancy in benefit to patients. Nonetheless, the concept of genetic medicine becoming true is the most exciting and victorious achievement of this millennium.

Various modalities of therapies that are based on the molecular etiology of the disorder are being researched; however, their detailed discussion is beyond the scope of this chapter. These molecular therapies can be broadly divided into the following:



1. *Gene therapy*: Where the defect in the genetic sequence is the target for correction and is achieved through bypassing the faulty sequence by “editing” the genome or by introducing a “transgene” or “silencing” the expression of that abnormal sequence. This is achieved by using viral or nonviral vectors. Gene therapy has been advancing through impressive techniques including the use of
  - a. viral vectors such as adeno-associated virus (AAV) to transfer the transgene to the tissue of interest;
  - b. nonviral vectors such as lipid nanomolecules;
  - c. genome editing techniques for correcting the missing or disrupted coding sequence [using zinc-finger nucleases, transcription activator-like effector nucleases, and the CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 system];
  - d. gene silencing or suppression by altering the expression of a faulty gene by “gene silencing or suppression” [Antisense oligonucleotide (ASO) and RNA interference (RNAi)].
2. *Protein modulation*: Where the protein encoded by the concerned gene is the target and pharmacological agents are used to modulate the function of that protein.
3. *Histone deacetylase (HDAC) inhibitors*: Used to modify histones given their importance in the epigenetic regulation of DNA structure and function.
4. *Novel uses of known pharmacological agents*: By exploiting their additional properties to treat genetic disorders.

Some examples of genetic disorders where above molecular-based therapies are being explored will be discussed in the following sections:

### 6.3.1 Gene therapy

#### **Primary immunodeficiencies**

Primary immunodeficiency diseases (PIDs) were one of the first genetic disorders attempted to be treated with gene therapy. These are a heterogeneous group of rare genetic diseases characterized by recurrent and life-threatening infections resulting from aberrations in development and differentiation of the immune system. While allogeneic hematopoietic stem cell (HSC) transplantation (HSCT) has been successfully used to treat severe PIDs with improving results, it remains an intensive procedure with a significant morbidity and mortality. Gene-therapy approaches were initially based on the use of gene transfer vectors derived from murine gammaretroviruses for patients affected with adenosine deaminase (ADA)-deficient severe combined immunodeficiency (SCID) who derived limited benefit from the genetic correction of either their peripheral blood lymphocytes or CD<sup>34+</sup> hematopoietic progenitors. Technical progress further led to the first unambiguous successful clinical application of gene therapy in patients affected with X-linked SCID (SCIDX-1), ADA–SCID, and Wiskott–Aldrich syndrome. Unfortunately, with the initial clear clinical benefits, the first serious complications of gene therapy also occurred. In a significant number of patients treated using murine gammaretroviral vectors, insertional oncogenesis events driven by the presence of the powerful viral enhancer elements resulted in acute leukemia, and fatal outcomes have been encountered in some cases.

HSCT and enzyme-replacement therapy (ERT) are the available forms of treatment for ADA–SCID, but each with drawbacks limiting their efficacies. Administration of mild myeloreductive chemotherapy with busulfan or melphalan and the withholding of ERT, as steps aimed at increasing the initial advantage of gene-corrected HSCs, were extremely effective in achieving immune reconstitution (increases in T-cell counts, normalization of T-cell function, and restoration of responses to vaccinations) in the majority of the treated patients who remained off of ERT. Interestingly, the immune recovery observed in ADA–SCID patients after gene therapy with gammaretroviral vectors occurred in the absence of insertional oncogenesis events, which distinguishes the experience in this disease from the other PIDs.

De Ravin et al. reported CRISPR/CAS9 system to repair a mutation in the CYBB gene of CD<sup>34+</sup> hemopoietic stem and progenitor cells (HSPCs) from patients with X-linked chronic granulomatous disease, which may avoid problems associated with gene therapy such as vector-related mutagenesis and dysregulated transgene expression [27,28].

#### **Duchenne muscular dystrophy**

Duchenne muscular dystrophy (DMD) is caused by mutations in the 2.4 mb dystrophin gene of which about 68% are intragenic large deletions, of which 66% are nested within a recombination prone, hotspot region spanning exons 45–55. The resulting joining of flanking exons by pre-mRNA splicing yields transcripts harboring

out-of-frame sequences and premature stop codons which are presumably degraded by nonsense-mediated mRNA decay [29].

Various therapies are being explored to modify the phenotype of this muscular dystrophy, which are as follows:

1. Mutation-specific exon skipping via modulation of pre-mRNA splicing by ASOs
2. Compensatory upregulation of utrophin (autosomal paralogue of dystrophin) by small molecule drugs or artificial transcription factors
3. Cell therapies involving allogenic myogenic stem/progenitor cell transplantation
4. Gene therapy based on shortened versions of dystrophin (microdystrophins)
5. Genome editing strategies based on sequence-specific programmable nucleases

In 2016 the US Food and Drug Administration (FDA)–approved Exondys 51 (Eteplirsen), the first exon skipping and exon inclusion drug, to treat patients with DMD. The exon skipping of DMD mRNA aims to restore the disrupted reading frame using ASOs, allowing the production of truncated but partly functional dystrophin proteins, and slow down the progression of the disease, converting to a milder Becker muscular dystrophy-like phenotype. Phase III study using Drisapersen for evaluating safety and efficacy in ambulatory boys above 5 years of age showed improvement in 6-m walk test and might be beneficial to less-impaired DMD patients. Clinical trials suggest that patients who benefit from exon 51 deletion have a more predictable phenotype than those who might benefit from exon 44 or 45, trials of which are under development [30,31].

### **Hemophilia B**

Hemophilia B is an X-linked recessive bleeding disorder due to deficiency of coagulation factor IX (FIX). While factor replacement by infusions as and when necessary has been the traditional treatment, gene therapy aims to ameliorate bleeding risk and provide endogenous FIX activity/synthesis through a single treatment, eliminating the requirement for FIX concentrate. A clinical trial with AAV-5 vector with a liver-specific promoter driving expression of a codon-optimized wild-type human FIX gene included 10 adults with hemophilia B (FIX = 2% of normal) and severe-bleeding phenotype. This trial reported a positive safety profile and improved stable FIX activity with a marked reduction in spontaneous bleeds and FIX concentrate use, without detectable cellular-immune responses against capsids [32]. Incorporation of a hyperactive gain of function R338L mutation (FIX Padua) in the FIX gene has also been reported to improve the overall efficacy, the use of transient immune suppression is speculated to control inflammatory immune response likely to be associated with hepatotoxicity [33].

### **Leber congenital amaurosis**

Leber congenital amaurosis is an autosomal recessive disorder with severe visual dysfunction with progressive retinal degeneration. Results of four human clinical trials have been reported to show safety and modest efficacy after subretinal injections of recombinant AAV 2 carrying the RPE65 gene. A 2-year follow up with such treatment was not associated with serious adverse events and improvement in visual function was observed in most patients, greater in younger patients with better baseline visual acuity [34].

### **Achondroplasia**

Achondroplasia is caused by heterozygous mutations in the fibroblast growth factor 3 (FGFR3) gene resulting in abnormal endochondral ossification leading to disproportionate rhizomelic short stature. FGFR3 signaling negatively regulates endochondral bone growth by inhibiting the rate of chondrocyte proliferation and the initiation of chondrocyte hypertrophy through activation of the Stat1 and MAPK pathways. Garcia et al. [35] reported the use of a FGFR3 decoy receptor (sFGFR3) to avoid the FGF ligand from binding to the receptor, thus preventing activation of the intracellular signaling directly downstream of mutant FGFR3 resulting in bone-growth activation. sFGFR3 was administered subcutaneously for 3 weeks to mice carrying the G380R mutation (Fgfr3ach/+ mice) at a dose of 2.5 mg/kg twice weekly from age 3 days to 3 weeks. This treatment was effective at restoring normal body, tail, and long bone lengths in treated Fgfr3ach/+ mice. There was reduction in respiratory failure with normal rib cage formation and decreased vertebral and skull deformities, which are common complications of achondroplasia. No obvious toxicity was observed, and the treatment did not affect reproduction. In addition the pelvic size showed improvement in primiparous-treated Fgfr3ach/+ females resulting in normal size litters. Thus the decoy approach could be viewed as a potential treatment for achondroplasia to restore the stature and to prevent complications due to the characteristic features of achondroplasia.

### ***Sickle cell disease***

The sickle cell allele is characterized by A to T substitution in the first exon of the beta globin gene and individuals homozygous for the mutation have classical sickle cell disease phenotype secondary to the loss of normal deformability of RBCs that adopt a sickle shape. There is a premature breakdown of RBCs, vascular occlusion leading to a cascade of hemolysis, ischemia, inflammation, and endothelial injury. Affected individuals experience pain crisis with ischemic episodes, susceptibility to infections, end organ injury, and early mortality although a variable clinical severity has been reported. At present the only curative therapy is allogenic stem cell transplantation along with hydroxyurea and supportive care.

Gene therapy is an attractive approach to managing sickle cell disease but would first demand safe and efficient gene transfer or correction of long-term repopulation of HSCs and a higher regulated stable gene expression. About three clinical trials have been reported, and one of them has used LentiGlobin BB305 vector expressing antisickling beta globin (T87Q) with 24% exogenous antisickling hemoglobin.

BCL11A is required for HbF expression, and erythroid-specific loss of BCL11A in a mouse model of sickle cell disease (SCD) is sufficient to reverse the hematopathological manifestations of SCD. Disruption of the erythroid-specific BCL11A enhancer and reduction of BCL11A expression below a critical threshold through RNAi are some of the approaches being considered. It is speculated that these newer gene engineering methods (compared to transduction of HSCs by lentiviral vectors) should facilitate the development of “second-generation” gene therapy approaches to SCD [36].

### ***Laminopathies***

Catelain et al. [37] reported stabilization and improvement of heart function after injecting bone morphogenic protein 2-committed embryonic stem cells and myoblasts at four sites on the anterior-lateral wall of the left ventricle of a mouse model of dilated cardiomyopathy caused by LMNA mutation (LmnaH222P/H222P mouse) after 4 and 8 weeks of transplantation. Maggi et al. [38] have discussed therapeutic possibilities of gene suppression by use of ASOs to skip LMNA exon 5 in treating specific laminopathies, while RNAi therapy has been tested in Hutchinson–Gilford progeria syndrome (HGPS), in which mutant LMNA mRNAs are selectively destroyed using short hairpin RNA or synthetic oligonucleotides with long half-lives. These approaches were shown to effectively restore several cellular and nuclear phenotypes.

### **6.3.2 The histone deacetylase inhibitors**

In addition to manipulating the genetic sequence or function, modification of the histone proteins associated with DNA is also of interest given their importance in the epigenetic regulation of DNA structure and function. HDACs remove the acetyl group from lysine residues within a range of proteins, including transcription factors and histones and their predominant function in deacetylation of lysine residues influencing expression of genes encoded by DNA linked to the histone molecule. HDAC inhibitors in turn regulate the activity of HDACs. HDAC6 expressions and activity are both increased in certain cancers, neurodegenerative diseases, and in Pkd1-mutant renal epithelial cells making HDAC inhibitors an attractive therapeutic drug [39].

### ***Kabuki syndrome***

KS is a multiple congenital anomalies syndrome characterized by characteristic facial features and varying degrees of mental retardation, caused by mutations in KMT2D/MLL2 and KDM6A/UTX genes. It is speculated that neurodevelopmental deficits in KS are maintained by an impairment of neurogenesis because of an imbalance between open and closed chromatin states for critical target genes. Bjornsson et al. [40] showed that memory deficits in a mouse model of KS can be prevented or even reversed through systemic delivery of a HDAC inhibitor that promotes open chromatin frame. Similar results with improved long-term memory have also been shown in Rubinstein–Taybi syndrome mouse models, caused due to mutation in another epigenome CREBBP.

### ***Autosomal dominant polycystic kidney disease***

Autosomal dominant polycystic kidney disease (ADPKD) is associated with progressive enlargement of multiple renal cysts, often leading to renal failure that cannot be prevented by a current treatment. Two proteins encoded by two genes are associated with ADPKD: PC1 (pkd1), primarily a signaling molecule, and PC2 (pkd2), a  $\text{Ca}^{2+}$  channel. Abnormal proliferation of epithelium lining the cysts and increased intracystic fluid secretion via the cystic fibrosis (CF) transmembrane conductance regulator (CFTR) are thought to contribute to cyst growth in

ADPKD. Dysregulation of cyclic adenosine monophosphate (cAMP) signaling is central to ADPKD with increased HDAC6 expression and activity in PKD1 mutant renal epithelial cells. HDAC6 inhibition reduces cell growth primarily by reducing intracellular cAMP and  $\text{Ca}^{2+}$  levels [41].

### 6.3.3 Protein modulation

Targeting the concerned protein instead of the gene itself is a novel way to tackle genetic disorders, and CF is a model for this approach. CF is an autosomal recessive genetic disorder that leads to chronic multisystem disease consisting of chronic sinopulmonary infections, malabsorption, and nutritional abnormalities. It is caused by mutations in CFTR gene that encodes the CFTR protein, a cyclic adenosine monophosphate-regulated ion channel. CFTR functions primarily as a chloride and bicarbonate channel and controls the movement of fluid into and out of epithelial cells lining the respiratory tract, biliary tree, intestines, vas deferens, sweat ducts, and pancreatic ducts.

Advances in the understanding of CF genetics have led to novel mutation-specific therapies. CFTR mutations can be broadly categorized into five classes based on the effect of the mutation on the protein. Class I mutations result in altered RNA processing and absent or truncated CFTR protein due to nonsense, splice site, or frameshift mutations. Class II mutations include F508del, leading to folding or maturation defects leading to little detectable CFTR in the cell membrane. Class III mutations, including the G551D mutation, are gating mutations that result in a nonfunctional CFTR protein with limited channel opening. Class IV and V mutations are associated with either reduced chloride conductance through the CFTR channel or reduced levels of the CFTR protein at the cell membrane, respectively. Class I–III mutations result in minimal protein function and are associated with the classic CF phenotype including pancreatic insufficiency. Individuals with at least one Class IV or V mutation (i.e., partial function mutations) typically have residual CFTR function and sufficient pancreatic function to absorb nutrients without supplemental pancreatic enzymes. However, certain mutations including F508del may cause problems at multiple steps and other modifier genes, and environmental factors may influence the variability of CF phenotype and treatment response.

CFTR protein modulators include “potentiators,” such as Ivacaftor, which work by opening the dysfunctional CFTR channel present at the cell surface (mutations Classes II and IV) and “correctors,” such as Lumacaftor, work by increasing trafficking of the CFTR protein to the cell membrane (Class III mutations). Compounds such as ataluren aim to mask abnormal gene sequence and allow the ribosome to “read-through” premature termination codons (PTCs) (Class I mutations) resulting in full-length functional protein [42].

### 6.3.4 Novel uses of known pharmacological agents

The discovery of newer molecular mechanisms has initiated the use of additional properties of some of the known drugs to treat genetic disorders. Many of these drugs are yet under clinical trials as their use is expected to be long term/lifelong, in dosages different from traditional indications and importantly in pediatric populations where they have not been used earlier.

#### *Joubert syndrome*

Srivastava et al. [43] reported that primary renal epithelial cells directly isolated from patient urine (human urine-derived renal epithelial cells) have elongated and disorganized primary cilia and that this ciliary phenotype is specifically associated with an absence of CEP290 protein. Treatment with the SHH-pathway agonist purmorphamine or cyclin-dependent kinase inhibition (using roscovitine and siRNA directed toward cyclin-dependent kinase 5) ameliorated the cilia phenotype. In addition, purmorphamine treatment was shown to reduce cyclin-dependent kinase 5 in patient cells, suggesting a convergence of these signaling pathways. These data implicate abnormal Hedgehog signaling as the cause of nephronophthisis in some ciliopathies and suggest that SHH agonists may be exploited therapeutically.

#### *Farber/Spinal muscular atrophy with progressive myoclonic epilepsy (SMA-PME)*

Acid ceramidase (*N*-acylsphingosine deacylase) degrades ceramid into sphingosine and free fatty acids within the lysosome and can also synthesize ceramid from sphingosine at neutral pH. Its deficiency is known to result in Farber lipogranulomatosis (Farber disease) and Spinal muscular atrophy with progressive myoclonic epilepsy (SMA–PME). The complementary DNA (cDNA) and *N*-acylsphingosine amidohydrolase (ASAH) gene encoding

acid ceramidase have been isolated, and recombinant enzyme is being considered for treatment of these disorders [44].

### ***Thanatophoric dysplasia and achondroplasia***

Jin et al. [45] reported reversal of the most severe, lethal form of FGFR3-related skeletal dysplasia (Thanatophoric dysplasia II) in mice by a 12-amino acid peptide called P3 that binds the extracellular domain of FGFR3 with high affinity decreasing the MAPK-signaling cascade. The neonatal TDII mice survived with partial rescue of endochondral bone growth.

Yamashita et al. [46] evaluated the use of statin to rescue bone growth in achondroplasia and thanatophoric dysplasia type I (TDI) as it stimulates bone growth through anabolic effects on chondrocytes. They used induced pluripotent stem cells generated from dermal fibroblasts of achondroplasia and TDI patients, which displayed decreased chondrogenic potential that was rescued by lovastatin treatment. In vivo, daily intraperitoneal injections of 1 mg/kg rosuvastatin from 3 to 6 weeks of age partially restored limb and skull lengths of *Fgfr3<sup>ach/+</sup>* mice showing potential therapeutic effects of statins. Meclozine, a known antihistamine drug used for motion sickness, has been evaluated by Matsushita et al. in transgenic *Fgfr3<sup>ach/+</sup>* mice that showed an increase in bone length with oral meclizine. However, the safety profile of chronic administration of statins and meclizine in children needs to be evaluated [47].

C-type natriuretic peptide (CNP) acts indirectly on FGFR3 signaling through the MAPK signaling pathway. Mice with achondroplasia treated with chronic CNP infusions were reported to have longer long bones and a larger foramen magnum. The initial phase-2 results of this trial demonstrate a favorable safety profile, and an increase in 50% of the annualized growth velocity following daily subcutaneous administration of 15 g/kg vosoritide for 12 months compared to the pretreatment growth velocity, although there was no improvement in their body proportions [48]. In a prospective observational study, Olney et al. [49] showed that CNP plasma levels are elevated in achondroplasia patients suggesting that these patients may have a natural resistance that could interfere with long term—treatment efficacy of vosoritide.

### ***Dystrophic epidermolysis bullosa***

Recessive dystrophic epidermolysis bullosa (RDEB) is a rare monogenic blistering disorder caused by the lack of functional type VII collagen, leading to skin fragility and subsequent trauma-induced separation of the epidermis from the underlying dermis. Nearly half of the patients with RDEB harbor at least one PTC mutation in COL7A1, which may be overcome by gentamicin-induced “read-through” and incorporation of an amino acid at the PTC site. In view of the aminoglycoside toxicity, Atanasova et al. [50] tested the FDA-approved drug amlexanox that was shown to increase COL7A1 transcript by inhibiting nonsense-mediated mRNA decay in cells from patients with RDEB demonstrating the potential of amlexanox for the treatment of patients with RDEB-harboring PTC mutation in COL7A1.

### ***Laminopathies***

Rapamycin is known to trigger lysosomal degradation of farnesylated prelamin A, the most toxic intermediate, and has been tested in cells from patients with Mandibuloacral dysostosis. A combination of all-trans retinoic acid and rapamycin was shown to significantly improve the cellular phenotype in HGPS, and the in vivo administration of temsirolimus, a rapamycin analog, was able to prevent the deterioration of cardiac function in hearts of mice with cardiomyopathy caused by the LMNA [38].

## **6.4 Conclusion/summary**

It is the need of the hour that clinical geneticists make a constant and conscious effort to understand molecular mechanisms leading to genetic disorders. The long-awaited genetic therapy will be based on molecular targets rather than on clinical clues alone. Precise description of phenotype, knowledge of changing phenotype with age, ethnicity, and contribution of environmental factors will strengthen genotype–phenotype correlation. Clinical geneticists who are traditionally the first ones to describe a phenotype and then order most relevant, focused genetic test might henceforth interpret results of a genetic test first and later confirm whether the given phenotype fits with the suspected molecular pathology. Genetic engineering and gene therapy are the future for managing genetic disorders and they are expected to correct the root cause unlike correcting symptoms only. There is a need and quest of making these therapies ideal—safer and longer lasting/lifetime correction with a “one-time” intervention.



## References

- [1] <https://www.sciencedirect.com/sdfe/pdf/download/eid/3-s2.0-B9780323054058001595/first-page-pdf> (accessed on 19.9.2019).
- [2] Shaheen R, Patel N, Shamseldin H, Alzahrani F, Al-Yamani R, AlMoisheer A, et al. Accelerated matchmaking of novel Dysmorphology syndromes through clinical and genomic characterization of a large cohort. *Genet Med* 2016;18(7):686–95.
- [3] Reardon W, Donnai D. Dysmorphology demystified. *Arch Dis Child Fetal Neonatal Ed* 2007;92(3):F225–9.
- [4] Kosilek RP, Frohner R, Würtz RP, Berr CM, Schopohl J, Reincke M, et al. Diagnostic use of facial image analysis software in endocrine and genetic disorders: review, current results and future perspectives. *Eur J Endocrinol* 2015;173(4):M39–44.
- [5] Allanson JE. Pitfalls of genetic diagnosis in the adolescent: the changing face. *Adolesc Med* 2002;13(2):257–68.
- [6] Biesecker L. Lumping and splitting. *Molecular biology in the genetics clinic. Clin Genet* 1998;3-7:s3.
- [7] Smedley D, Robinson PN. Phenotype driven strategies for exome prioritization of human Mendelian disease genes. *Genome Med*. 2015;7(1):81. Available from: <https://doi.org/10.1186/s13073-015-0199-2>, eCollection 2015.
- [8] Deelan P, van Dam S, Herkert JC, Karjalainen JM, Brge H, Abbott KM. Improving the diagnostic yield of exome-sequencing by predicting gene-phenotype associations using large-scale gene expression analysis. *Nat Commun*. 2019;10(1):2837. Available from: <https://doi.org/10.1038/s41467-019-10649-4>.
- [9] Joshi C, Kolbe AL, Mansilla MA, Mason AO, Smith RJH, Campbell CA. Reducing the cost of the diagnostic odyssey in early onset epileptic encephalopathies. *BioMed Res Int* 2016;2016:6421039.
- [10] Petryk A. Holoprosencephaly: signaling interactions between the brain and the face, the environment and the genes and the phenotypic variability in animal models and humans. *Rev Dev Biol* 2015;4(1):17–32.
- [11] Hong M, Krauss RS. Ethanol itself is a holoprosencephaly-inducing teratogen. *PLoS One* 2017;12(4):e0167440.
- [12] Pineda-Alvarez DE, Dubourg C, David V, Roessler E, Muenke M. Current recommendations for the molecular evaluation of newly diagnosed holoprosencephaly patients. *Am J Med Genet C Semin Med Genet* 2010;154C(1):93–101 5.
- [13] Sowiriska-Seidler A, Socha M, Jamsheer A. Split-hand/foot malformation-molecular cause and implications in genetic counseling. *J Appl Genet* 2014;55:105–15.
- [14] Klopocki E, Lohan S, Doelken SC, Stricker S, Ockeloen CW, Soares Thiele de Aguiar R, et al. Duplications of BHLHA9 are associated with ectrodactyly and tibia hemimelia inherited in non-Mendelian fashion. *J Med Genet* 2012;49(2):119–25.
- [15] Altmüller F, Lissewski C, Bertola D, Flex E, Stark Z, Spranger S, et al. Genotype and phenotype spectrum of NRAS germline variants. *Eur J Hum Genet* 2017;25(7):823–31.
- [16] Santoro C, Pacileo G, Limongelli G, Scianguetta S, Giugliano T, Piluso G, et al. LEOPARD syndrome: clinical dilemmas in differential diagnosis of RASopathies. *BMC Med Genet* 2014;15:44.
- [17] Gripp KW, Aldinger KA, Bennett JT, Baker L, Tusi J, Powell-Hamilton N, et al. A novel rasopathy caused by recurrent de novo missense mutations in PPP1CB closely resembles Noonan syndrome with loose anagen hair. *Am J Med Genet A* 2016;170(9):2237–47.
- [18] Wormana HJ, Bonne G. “Laminopathies:” a wide spectrum of human diseases. *Exp Cell Res* 2007;313(10):2121–33.
- [19] Madhivanan K, Aguilar RC. Ciliopathies: the trafficking connection. *Traffic* 2014;15(10):1031–56.
- [20] Shaheen R, Szymanska K, Basu B, Patel N, Ewida N, Faqeih E, et al. Characterizing the morbid genome of ciliopathies. *Genome Biol* 2016;17(1):242.
- [21] Zhou AX, Hartwig JH, Akyürek LM. Filamins in cell signaling, transcription and organ development. *Trends Cell Biol* 2010;20(2):113–23.
- [22] Fürst DO, Goldfarb LG, Kley RA, Vorgerd M, Olivé M, van der Ven PF. Filamin C-related myopathies: pathology and mechanisms. *Acta Neuropathol* 2013;125(1):33–46.
- [23] Izumi K. Disorders of transcriptional regulation: an emerging category of multiple malformation syndromes. *Mol Syndromol* 2016;7:262–73.
- [24] Larizza L, Finelli P. Developmental disorders with intellectual disability driven by chromatin dysregulation: clinical overlaps and molecular mechanisms. *Clin Genet* 2019;. Available from: <https://doi.org/10.1111/cge.13365> [Epub ahead of print, accessed on August 23rd 2018].
- [25] Butcher DT, Cytrynbaum C, Turinsky A, Siu MT, Inbar-Feigenberg M, Mendoza-Londono R, et al. CHARGE and Kabuki syndromes: Gene-specific DNA methylation signatures identify epigenetic mechanisms linking these clinically overlapping condition. *Am J Hum Genet* 2017;100:773–88.
- [26] Osborne RJ, Thornton CA. RNA-dominant diseases. *Hum Mol Genet* 2006;15(2):R162–9.
- [27] Candotti F. Advances of gene therapy for primary immunodeficiencies [version 1; referees: 2 approved]. *F1000Research* 2016;5(F1000 Faculty Rev):310.
- [28] De Ravin SS, Li L, Wu X, Choi U, Allen C, Koontz S, et al. CRISPR-Cas9 gene repair of hematopoietic stem cells from patients with X-linked chronic granulomatous disease. *Sci Transl Med* 2017;9(372) pii: eaah3480.
- [29] Maggio I, Chen X, Goncalves MAFV. The emerging role of viral vectors as vehicles for DMD gene editing. *Genome Med* 2016;8:59.
- [30] Goemans N, Mercuri E, Belousova E, Komaki H, Dubrovsky A, McDonald CM, et al. A randomized placebo-controlled phase 3 trial of an antisense oligonucleotide drisapersen, in Duchenne muscular dystrophy. *Neuromuscul Disord* 2017;30 pii: S0960-8966(17)30200-6.
- [31] Antony K, Arechavala-Gomez V, Ricotti S, Feng L, Junghra N, Tasca G, et al. Biochemical characterization of patients with in frame or out of frame DMD deletions pertinent to exon 44 or 45 skipping. *JAMA Neurol* 2014;71(1):32–40.
- [32] Miesbach W, Meijer K, Coppens M, Kampmann P, Klamroth R, Schutgens R, et al. Gene therapy with adeno-associated virus vector 5-human factor IX in adults with hemophilia B. *Blood* 2018;131(9):1022–31 pii: blood-2017-09-804419.
- [33] Evens H, Chuah MK, Venden Driessche T. Hemophilia gene therapy: from trailblazer to gamechanger. *Hemophilia* 2018;24 (suppl6):50–9.
- [34] Weleber RG, Pennesi ME, Wilson DJ, Kaushal S, Erker LR, Jensen L, et al. Results at 2 years after gene therapy for RPE65-deficient Leber congenital amaurosis and severe early-childhood-onset retinal dystrophy. *Ophthalmology* 2016;123(7):1606–20.
- [35] Garcia S, Dirat B, Tognacci T, Rochet N, Mouska X, Bonnafous S, et al. Postnatal soluble FGFR3 therapy rescues achondroplasia symptoms and restores bone growth in mice. *Sci Transl Med* 2013;5:203ra124.



- [36] Hoban MD, Orkin AH, Bauer DE. Genetic treatment of a molecular disorder: gene therapy approaches to sickle cell disease. *Blood* 2016;127(7):839–48.
- [37] Catelain C, Riveron S, Papadopoulos A, Mougenot N, Jacquet A, Vauchez K, et al. Myoblasts and embryonic stem cells differentially engraft in a mouse model of genetic dilated cardiomyopathy. *Mol Ther* 2013;21(5):1064–75.
- [38] Maggi L, Carboni N, Bernasconi P. Skeletal muscle laminopathies: a review of clinical and molecular features. *Cells* 2016;5(3) pii: E33.
- [39] Bassett SA, Barnett MP. The role of dietary histone deacetylases (HDACs) inhibitors in health and disease. *Nutrients* 2014;6(10):4273–301.
- [40] Bjornsson HT, Benjamin JS, Zhang L, Weissman J, Gerber EE, Chen YC, et al. Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of Kabuki syndrome. *Sci Transl Med* 2014;6(256):256ra135.
- [41] Yanda MK, Liu Q, Cebotaru V, Guggino WB, Cebotaru L. Histone deacetylase 6 inhibition reduces cysts by decreasing cAMP and  $\text{Ca}^{2+}$  in knock-out mouse models of polycystic kidney disease. *J Biol Chem* 2017;292(43):17897–908.
- [42] Martiniano SL, Sagel SD, Zemanick ET. Cystic fibrosis: a model system for precision medicine. *Curr Opin Pediatr* 2016;28(3):312–17.
- [43] Srivastava S, Ramsbottom SA, Molinari E, Alkanderi S, Filby A, White K, et al. A human patient-derived cellular model of Joubert syndrome reveals ciliary defects which can be rescued with targeted therapies. *Hum Mol Genet* 2017;26(23):4657–67.
- [44] Frohbergh M, He X, Schuchman EH. The molecular medicine of acid ceramidase. *Biol Chem* 2015;396(6-7):759–65.
- [45] Jin M, Yu Y, Qi H, Xie Y, Su N, Wang X, et al. A novel FGFR3-binding peptide inhibits FGFR3 signaling and reverses the lethal phenotype of mice mimicking human thanatophoric dysplasia. *Hum Mol Genet* 2012;21:5443–55.
- [46] Yamashita A, Morioka M, Kishi H, Kimura T, Yahara Y, Okada M, et al. Statin treatment rescues FGFR3 skeletal dysplasia phenotypes. *Nature* 2014;513:507–11.
- [47] Matsushita M, Kitoh H, Ohkawara B, Mishima K, Kaneko H, Ito M, et al. Meclozine facilitates proliferation and differentiation of chondrocytes by attenuating abnormally activated FGFR3 signaling in achondroplasia. *PLoS One* 2013;8:e81569.
- [48] Noonberg S. BMN 111: vosoritide for achondroplasia. *Biomarin R&D Day* 2016;98–135:70.
- [49] Olney RC, Prickett TC, Espiner EA, Mackenzie WG, Duker AL, Ditro C, et al. C-type natriuretic peptide plasma levels are elevated in subjects with achondroplasia, hypochondroplasia, and thanatophoric dysplasia. *J Clin Endocrinol Metab* 2015;100:E355–9.
- [50] Atanasova VS, Jiang Q, Prisco M, Gruber C, Piñón Hofbauer J, Chen M, et al. Amlexanox enhances premature termination codon read-through in COL7A1 and expression of full length type VII collagen: potential therapy for recessive dystrophic epidermolysis bullosa. *J Invest Dermatol* 2017;137(9):1842–9.



# Disorders of sex development

*Pierre Sinajon<sup>1,2,3</sup>, Saleh Albanyan<sup>3</sup> and David Chitayat<sup>1,3</sup>*

<sup>1</sup>The Prenatal Diagnosis and Medical Genetics Program, Department of Obstetrics and Gynecology, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada <sup>2</sup>Department of Pediatrics, Division of Genetics, University of Rochester Medical Center, NY, United States <sup>3</sup>Division of Clinical and Metabolic Genetics, Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada

## 7.1 Introduction

Normal sex development involves a cascade of many autosomal and X-linked activating and repressing genes functioning along a short period of time in a precise sequence. The process of sex determination includes the following four layers:

1. Determination of the chromosome sex
2. Determination of the gonadal sex
3. Development of the internal genitalia
4. Development of the external genitalia

In this chapter, we will not focus on issues of psychosocial sex determination.

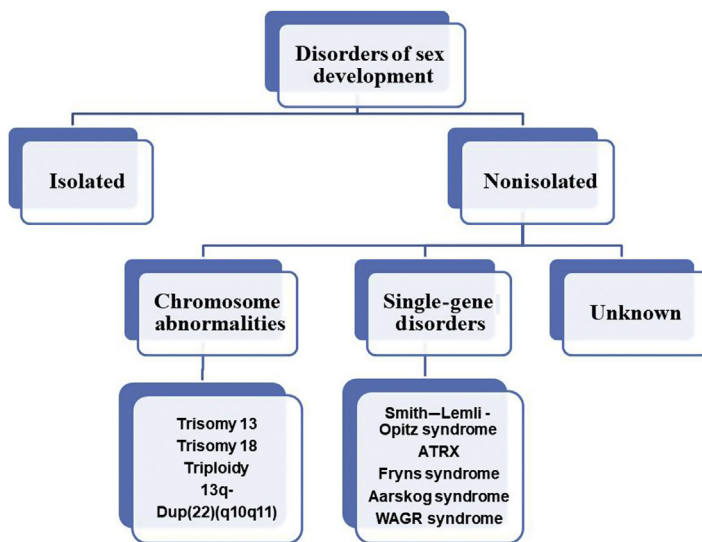
Disorders of sex determination (DSDs) consist of a heterogeneous group of congenital conditions associated with atypical development of the internal and external genitalia with an estimated incidence of 1:200–1:4500 [1].

The nomenclature used in the past for the description of genital abnormalities was descriptive, offending and did not use our current knowledge of karyotype and the molecular etiology of these conditions. Thus a new nomenclature was introduced at the Chicago Consensus Conference 2005 [2], which defined the DSD as a congenital condition in which the development of chromosomal, gonadal, or anatomical sex is atypical.

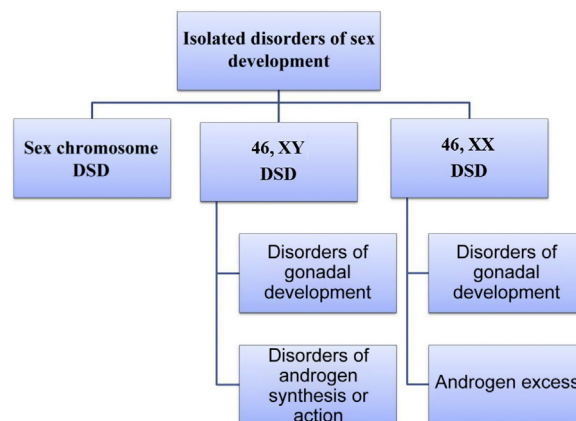
DSD can be isolated or associated with other major abnormalities (Fig. 7.1). When associated with other abnormalities, they are usually the result of a variety of chromosome abnormalities, autosomal or X-linked single-gene disorders, which affects different parts of the normal cascade (Fig. 7.2). This chapter will focus on isolated/non-syndromic disorders of sex determination.

Normally, it is the presence of the SRY gene on the Y chromosome that starts the cascade of events that results in the testicular gonadal determination in males. In males the secretion of testicular hormones and the function of their receptors determine the normal development of the male internal and external genitalia. The lack of the SRY gene as well as the activity of many genes will determine the ovarian gonadal development. The lack of testicular hormones and masculinizing hormones in general will determine the normal female internal and external genitalia (Figs. 7.3 and 7.4).

In humans the gonads are populated by primordial germ cells, deriving from the yolk sac wall early in week 5 postconception. It was thought that sex differentiation into testis required the expression of specific genes, while their absence would result in the development of an ovary. However, it is now clear that gonadal sex is determined by antagonistic interactions between ovarian and testicular pathways [3,4] around week 6. In the female embryo, germ cells are exposed to high levels of retinoic acid, which induces the expression of *STRA8*, leading to germ cell meiosis and development of oocytes. In the developing testis the absence of retinoic acid causes the germ cells to develop into gonocytes, which differentiate into spermatogonia and proliferate by mitosis. Meiosis happens only after puberty [5].



**FIGURE 7.1** Disorders of sex development—initial assessment.



**FIGURE 7.2** Differential diagnosis in isolated disorders of sex development.

There are three main crucial differences between ovarian and testicular function (Fig. 7.3):

1. The testes produce testosterone early in embryogenesis to induce the male external genital development, while the ovaries do not produce hormones until puberty.
2. Oogenesis starts and ends prenatally, while the testes start spermatogenesis only at puberty.
3. The ovary functions hormonally as long as the process of follicle development and/or maintenance exists, while the testes continue producing testosterone even in the absence of spermatogenesis.

The management of patients with DSD needs to be individualized to the patient affected taking into account the parental and the families' psychosocial environment. This includes decisions in regard to sex of rearing, surgical interventions, hormone treatment, and potential for fertility preservation. This is an urgent clinical situation associated with great distress for the family. It requires immediate assessment and counseling, and if possible, involving a multidisciplinary team with an expertise in DSD including endocrinology, clinical genetics, genetic counseling, urology, obstetrics and gynecology, social work, and psychology/psychiatry.

## 7.2 Sex chromosome disorder of sex development

Most sex chromosome abnormalities are not associated with genital abnormalities. Turner syndrome is the most common chromosome abnormality at conception with 2% of embryos being affected with the condition.

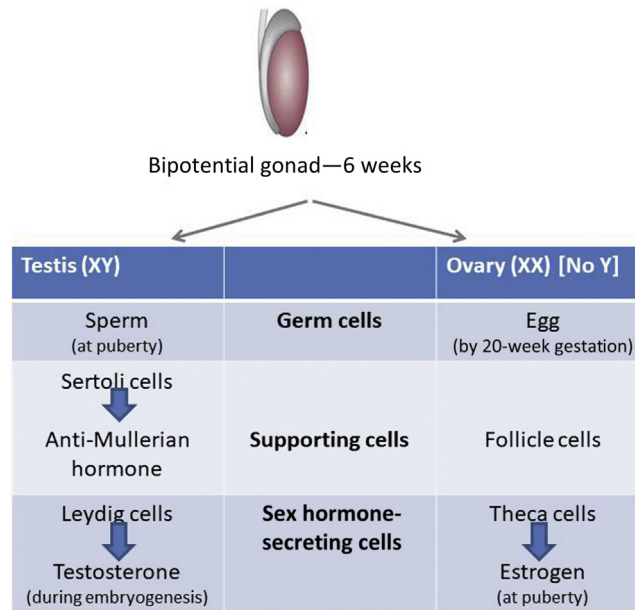


FIGURE 7.3 Gonadal differentiation and function.

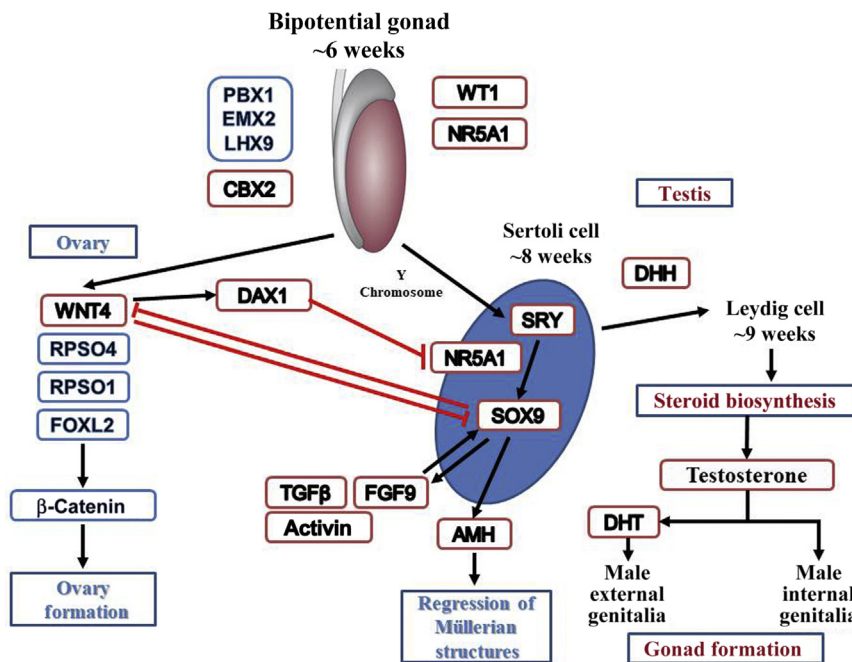


FIGURE 7.4 Diagram of the known genes involved in testicular and ovarian differentiation and function. Source: Adapted from Ahmed SF, Bashamboo A, Lucas-Herald A, McElreavey K. Understanding the genetic aetiology in patients with XY DSD. *Br Med Bull* 2013;106:67–89; Rey R, Josso N, Racine C. Sexual differentiation. In: De Groot LJ, Chrousos G, Dungan K, et al., editors. *Endotext*. South Dartmouth, MA: MDText.com, Inc.; 2000. Available from: <<https://www.ncbi.nlm.nih.gov/books/NBK279001/>>.

The condition carries a high lethality rate, and the incidence among newborn females is only 1:2500. The condition is associated with 45, X karyotype in 60%, mosaicism (45, X/46, XX; 45, X/46, XY; 45, X/47, XXX; and 45, X/46, XX/47, XXX) in 15%, structural abnormalities involving the X chromosome [46, X, i(Xq), 46, X, Xp- and 46, X, r(X)] in 10%, mosaicism with structural X chromosome abnormality in 10% and others in 5% [6]. The condition is characterized prenatally by a cystic hygroma or increased nuchal translucency as well as hydrops fetalis and left-sided heart lesions. Postnatally, girls with this condition have characteristic facial features with epicanthal folds, downslanting palpebral fissures, low-set and prominent ears, and micrognathia. Other findings include short and webbed neck with a low nuchal hairline, lymphedema of hands and feet, along with left-sided cardiac

anomalies in 50% of the cases (hypoplastic left heart, coarctation of the aorta, bicuspid aortic valve, and aortic stenosis). It is also associated with an increased risk for aortic dissection and horseshoe kidneys. Other typical presentations include short stature and no spontaneous puberty.

Genetic analysis for Y chromosome material is important because these girls have an increased risk for gonadoblastoma, dysgerminoma, and aortic dissection. This can be detected by microarray analysis. A large series of British women with Turner syndrome reported a decreased incidence of breast cancer but an increased risk for gonadoblastoma, corpus uteri cancer, and possibly childhood brain cancers [7–9].

47, XXY (Klinefelter syndrome) has an incidence of about 1 in 500 males. Affected boys have no facial dysmorphism and normal external genitalia. They may present with tall stature, small testes, delayed puberty, gynecomastia, and infertility at puberty or after puberty. Boys with Klinefelter syndrome often manifest behavior difficulties, dyslexia, defects in executive function, and autism spectrum disorders. An improved outcome has been reported in association with early treatment with androgens. Standard treatment for Klinefelter syndrome consists of androgen replacement therapy in adolescence to offset testosterone deficiency. Such treatment has a beneficial effect on the physical and behavioral manifestations of these patients. Recently, early androgen supplementation in children with this condition has been found to improve behavioral functioning. The optimal timing of hormonal therapy has not been established, and further studies are required before implementing this treatment as a standard of care [10].

45, X/46, XY previously known as mixed gonadal dysgenesis is a rare condition with an estimated incidence rate of less than 1/15,000 live births in Denmark [6]. The presence of the Y chromosome material in the gonadal ridge has been suggested as the determining factor for the early testicular gonadal differentiation [11–13] and thus represents a wide spectrum of phenotypes; from Turner syndrome females to phenotypically normal males with genital abnormalities and some with short stature. Gonadal function in most 45, X/46, XY males, even those with genital ambiguity, seems sufficient for spontaneous puberty and patients appear to be benefitted from GH treatment [14].

### 7.3 46, XY disorders of sexual differentiation

46, XY DSD can be divided broadly into two categories; disorders of testicular (gonadal) development and disorders of androgen synthesis and impaired testosterone action all leading to feminization of the genitalia.

#### 7.3.1 Disorders of testicular (gonadal) development

Disorders of testicular development are characterized by absent or small testes on ultrasound and/or palpation and the presence of Müllerian structures (including the uterus and fallopian tubes) with varying degrees of feminization of the external genitalia as well as decreased blood levels of testosterone, dehydroepiandrosterone, and androstenedione [15].

At approximately 6 weeks from conception the gonads are undifferentiated and can form either ovaries or testis [16]. Multiple genes are involved in testicular differentiation (Fig. 7.4), and both haploinsufficiency with loss-of-function mutations and duplication with gain-of-function mutations are known to be associated with 46, XY gonadal DSD. The *SRY* gene, on the distal short arm of Y [Yp11.2], encodes a transcription factor, which starts the cascade that directs the bipotential gonad to form testis [17]. Mutations/Deletions of the *SRY* gene result in complete gonadal dysgenesis or 46, XY pure gonadal dysgenesis and have been identified in approximately 15% of individuals with Swyer syndrome. It is also known as 46, XY complete gonadal dysgenesis or 46, XY pure gonadal dysgenesis, which results in female internal and external genitalia.

The *SRY* with the *NR5A1* activates *SOX9* and anti-Müllerian hormone (AMH) formation by the Sertoli cells [18]. Mutations in the *SOX9* gene (OMIM 608160) present with campomelic dysplasia as well as sex reversal in 46, XY. A heterozygous deletion of approximately 240-kb between 405 and 645 kb upstream of the *SOX9* transcription start site was reported in association with 46, XY with a normal external female phenotype/severe ambiguous and asymmetric external genitalia [19]. A heterozygous duplication upstream of the *SOX9* gene on chromosome 17 resulted in a 46, XX male [20].

The ligand *FGF9* and the signaling molecule *WNT4* are expressed in the undifferentiated gonad and continue to be expressed in the ovary as well as in the testis and provide opposing signals that determine gonadal differentiation. The *FGF9* promotes testicular differentiation, and the Wnt family member (*WNT*)/R-spondin 1



signaling and forkhead box L2 (*FOXL2*) drive female sex determination in XX gonads, and promotes ovarian differentiation. The *SRY* gene initiates a feed-forward loop that interacts with *SOX9* and *FGF9* resulting in the upregulation of *FGF9* and represses *WNT4* and thus establishes testicular formation. The receptor for the *FGF9* in the developing testis is *FGFR2* [21]. Another important gene for gonadal development at the bipotential gonad level is the *NR5A1* gene (also known as *SF-1*) (OMIM 184757). This is an orphan nuclear receptor that is expressed in the hypothalamus, pituitary, gonads, and adrenal glands [16]. Pathogenic variants in this gene result in some females (but not all) presenting with premature ovarian failure (POF) and in a variety of DSD in 46, XY ranging from ambiguous genitalia to female external genitalia with complete to incomplete regression of the Müllerian duct derivatives [22–24]. The mutations p.G35D, p.G35E, p.R92Q, and p.R255L also result in adrenal insufficiency combined with gonadal dysfunction. *DAX1* is another orphan nuclear receptor with roles in the hypothalamus, pituitary, gonads, and adrenal glands [25].

The Wilms tumor factor 1 (*WT1*) mutations present with an increased risk for nephroblastoma and genitourinary abnormalities. Two conditions commonly described are Denys–Drash and Frasier syndromes. Denys–Drash syndrome presents with the triad of genitourinary abnormalities, renal impairment, and Wilms tumor [26]. Frasier syndrome presents with later onset renal impairment (focal segmental glomerulonephritis) and a risk for Wilms tumor, though not as high compared to other *WT1* related conditions. Other genes of interest in 46, XY gonadal DSD include *CBX2* and *DHH*. *CBX2* is involved in the regulation of homeotic genes and of the bipotential gonad [16]. Mutations in the *CBX2* gene have been described in 46, XY DSD presenting with normal female internal and external genitalia. These studies suggest that *CBX2* could be responsible for the repression of ovarian development [27]. In XY individuals, *MAP3K1* mutations appear to shift signaling pathways to suppress *SOX9* and promote ovarian differentiation [28]. *DHH* is involved in Leydig cell differentiation, which occurs at approximately 9 weeks of embryonic development [16]. Mutations in *DHH* cause complete or partial gonadal dysgenesis along with minifascicular neuropathy (OMIM 605423) [29]. Mutations in *DAX1* can cause congenital adrenal hypoplasia and hypogonadotropic hypogonadism (OMIM 300473) and complete or partial gonadal development with ambiguous external genitalia [30]. Other genes associated with XY gonadal dysgenesis include *DMRT1*, *DMRT2*, *SOX3*, and *SOX8* among others [31].

### 7.3.2 Disorders of androgen synthesis

Disorders of androgen synthesis are conditions characterized by a lack of Müllerian structures due to AMH production by the testes in a 46, XY with genital abnormalities to normal female external genitalia. Disorders of androgen production can be divided into two categories: those associated with congenital adrenal hyperplasia (CAH) and those associated with normal adrenal function. So far seven enzymes involved in the production of testosterone and dihydrotestosterone (DHT), responsible for 46, XY disorders of sex differentiation, have been identified (Fig. 7.5).

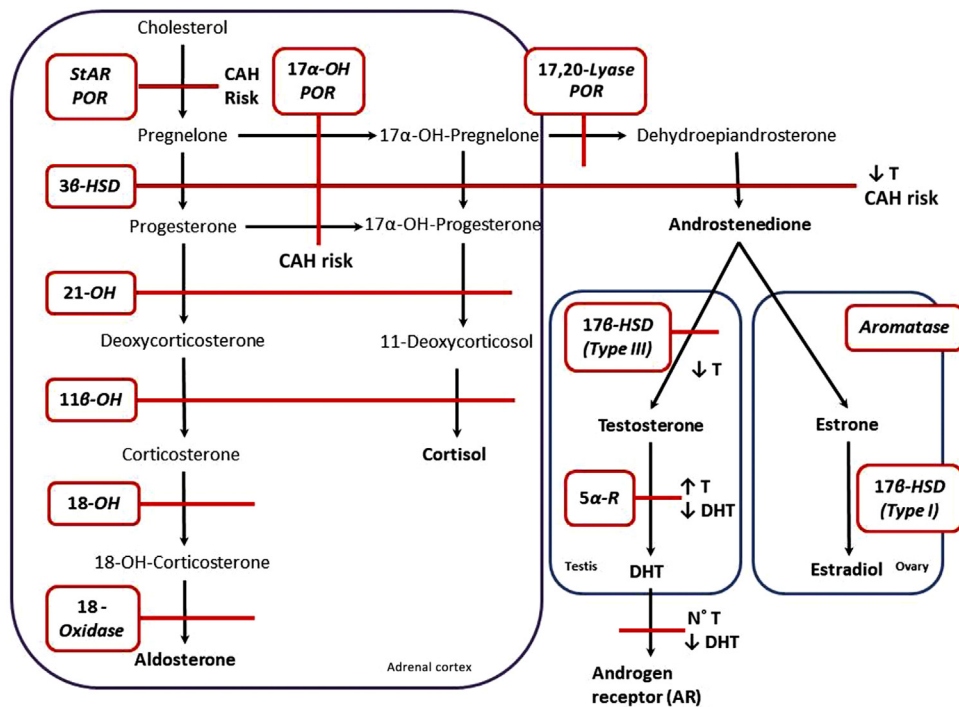
#### 7.3.2.1 Disorders of androgen synthesis associated with adrenal dysfunction

Defects early in the pathway result in CAH with adrenal insufficiency. These enzymes are present in both the adrenal cortex and the gonads.

The first enzymatic step involves the cleavage of cholesterol to pregnenolone through steroidogenic acute regulatory (*StAR*) protein and p450 oxidoreductase (*POR*). Defects in *StAR* cause lipid CAH (OMIM 201710), an autosomal recessive condition characterized by lipid accumulation, severe salt wasting and genital ambiguity ranging from hypospadias to complete female external genitalia. *POR* defects are more commonly associated with Antley–Bixler syndrome with genital anomalies and disordered steroidogenesis (OMIM 201750) but have been associated with a nonsyndromic presentation (OMIM 613571), which cause a similar picture to *StAR* mutations but the absence of CAH [32].

The next major step is through 3 $\beta$ -hydroxysteroid dehydrogenase (3 $\beta$ -HSD), which converts pregnenolone to progesterone, 17 $\alpha$ -OH pregnenolone to 17 $\alpha$ -OH progesterone, and DHEA to androstenedione. Defects in 3 $\beta$ -HSD are at increased risk for salt wasting and will have decreased testosterone along with ambiguous genitalia characterized by hypospadias, micropenis, and bifid scrotum (OMIM 201810).

The last enzyme with an increased risk for adrenal insufficiency and ambiguous genitalia is 17 $\alpha$ -hydroxylase, which converts pregnenolone and progesterone to their 17 $\alpha$  hydroxylated forms. There is clinical overlap with a combined deficiency in 17 $\alpha$ -hydroxylase and 17,20-lyase. Isolated 17,20-lyase deficiency has normal



**FIGURE 7.5** Steroid hormone synthesis pathway with biochemical abnormalities in 46, XY DSD.

11 $\beta$ -OH, 11 $\beta$ -Hydroxylase; 17 $\alpha$ -OH, 17 $\alpha$ -hydroxylase; 17 $\beta$ -HSD, 17 $\beta$ -hydroxysteroid dehydrogenase; 18-OH, 18-hydroxylase; 21-OH, 21-hydroxylase; 3 $\beta$ -HSD, 3 $\beta$ -hydroxysteroid dehydrogenase; 5 $\alpha$ -R, 5 $\alpha$ -reductase; CAH, congenital adrenal hyperplasia; DSD, disorder of sex development; POR, P450 oxidoreductase; StAR, steroid acute regulatory protein. Source: Adapted from Wherrett, DK. Approach to the infant with a suspected disorder of sex development. *Pediatr Clin N Am* 2015;62(4):983–99.

adrenal function and variable abnormalities of male phenotype since the 17,20-lyase is present in the gonads only.

### 7.3.2.2 Disorders of androgen synthesis associated without adrenal dysfunction

Within the gonads the conversion of androstenedione into testosterone is through 17 $\beta$ -hydroxysteroid dehydrogenase type III. Mutations in this gene cause an autosomal recessive disorder (OMIM 264300) with female external genitalia, male gonadal derivatives, absent Müllerian structures, and infertility in 46, XY individuals [33]. Biochemically they present with decreased testosterone. These patients are sometimes difficult to distinguish from 5 $\alpha$ -reductase deficiency and partial androgen insensitivity syndrome (AIS) [34]. Further testing with ACTH stimulation or hCG stimulation may be needed [15] to confirm the diagnosis although DNA analysis is probably easier to perform.

Testosterone and DHT are the end products for testis steroid hormone synthesis. Mutations occur in the 5 $\alpha$ -reductase gene (*SRD5A2*), which catalyzes the formation of DHT from testosterone, associated with low levels of DHT, normal/increased levels of testosterone, and high testosterone/DHT levels. Patients have normal male internal genitalia but undermasculinization of the external genitalia due to low levels of DHT [34]. The external genitalia range from a female phenotype to a small phallus with severe hypospadias. These patients may have normal sperm production.

### 7.3.3 Disorders of androgen response

The most common cause of 46, XY disorders of sex development is the AIS (OMIM 300068) where the androgen receptor (AR) is unable to activate due to the inability of testosterone or DHT to bind to the receptor [35]. This leads to a lack of effect of androgens on the genital development. These conditions are inherited in an X-linked manner with a wide range in phenotypes. Complete androgen insensitivity syndrome (CAIS) has an estimated prevalence of at least 1:99,000 [36] presenting with normal female genitalia and blind ending vaginal pouch. Partial androgen insensitivity syndrome (PAIS) occurs when there is residual AR function. PAIS has an

estimated prevalence of 1:8000, and hypospadias is the common finding. Mild androgen insensitivity syndrome (MAIS) is the least severe in the spectrum of AIS. MAIS usually presents with no genital abnormalities but can be suspected in the context of pubertal gynecomastia or unexplained infertility [33]. AR mutations can be located outside the coding region and associated with AIS [37,38]. In some instances, no AR mutations have been detected, suggesting that other proteins located beyond AR influence testosterone signaling [39].

## 7.4 46, XX disorders of sex development

46, XX disorders of sex development can be the result of abnormal gonadal (ovarian) development or excess androgen level due to abnormal synthesis or androgen exposure.

### 7.4.1 Ovarian development

There are two types of XX sex reversal, which are as follows:

1. Patients with SRY positive caused by translocation of the SRY gene to another chromosome, usually the X chromosome or, rarely, an autosome
2. Patients with SRY-negative XX males

Loss-of-function mutations in genes coding for ovarian formation and function are associated with ovarian dysgenesis and/or accelerated loss of primordial follicles, causing POF/premature menopause.

An important gene in the differentiation of the bipotential gonad into ovaries is the *WNT4* gene. The *WNT4* gene is a member of the WNT family of secreted molecules that function in a paracrine manner. The WNT proteins are ligands to members of the frizzled (FZ) family of cell surface receptors and possibly to the single-pass transmembrane protein LDL-receptor-related proteins 5 and 6 (*LRP5* and *LRP6*) [40]. The binding of WNT to FZ leads to reduced degradation of  $\beta$ -catenin with consequent  $\beta$ -catenin-dependent activation of T-cell factor/lymphocyte-enhancer factor transcription factors, and induction of WNT-responsive genes [41]. *WNT4* is produced in ovarian pregranulosa cells and *WNT4* upregulates *DAX1* [42], a gene known to antagonizes *NR5A1*, and thus inhibiting steroidogenic enzymes. *WNT4*-knockout XX mice were masculinized, had no Müllerian ducts derivatives but Wolffian ducts with expression of the steroidogenic enzymes  $3\beta$ -OH dehydrogenase and  $17\alpha$ -OH, crucial in the production of testosterone and thus normally suppressed in the developing female ovary. The ovaries of these mice also had decreased number of oocytes, showing the importance of *WNT4* in maintaining the female germ cells and thus normal ovarian function [43] (unlike testicular function that continues in the absence of sperm).

In humans, more copies of *WNT4*, caused by duplication of chromosome 1p31–p35, is known to be associated with male-to-female sex reversal exhibiting ambiguous genitalia, severe hypospadias, streak gonads, and remnants of both Müllerian and Wolffian ducts [42]. However, homozygotes with pathogenic variants in the *WNT4* gene result in SERKAL syndrome [44] presenting with 46, XX with female-to-male sex reversal, ambiguous genitalia, gonadal morphology ranging from ovotestis to normal testis, renal agenesis, adrenal hypoplasia, pulmonary abnormalities, and cardiac abnormalities. Mutations were also identified in women with absent Müllerian structures and clinical signs of androgen excess as well as women with findings resembling Mayer–Rokitansky–Kuster–Hauser syndrome [45].

Another important gene in the formation and function of the ovaries is the *FOXL2*. Pathogenic variants in this gene result in blepharophimosis, ptosis, and epicanthus inversus syndrome (BPES) (OMIM 110100). In BPES I the condition is associated with POF and in BPES type II it is not associated with POF.

In mice the continued expression of *FOXL2* in the ovary is essential in maintaining the ovarian function. Loss of this gene expression reprograms granulosa and theca cells into cells that are similar to Sertoli and Leydig cells, respectively [30,46–49]. *FOXL2* also has a role in stimulating the expression of the gonadotropin-releasing hormone receptor. The *FOXL2* expression precedes glycoprotein hormone, which is a common subunit to FSH, LH, and TSH expression in the pituitary gland [50].

Another important gene involved in ovarian development is the *RSPO1* (roof plate-specific spondin-1). This is a key regulator of female sex differentiation, which acts as an activator of the canonical WNT/ $\beta$ -catenin pathway opposing testis formation, with *WNT4* as the key ligand [51,52]. During sex differentiation, significantly higher

expression of *RSPO1* was detected in ovary compared to testis [52], supporting an important role in female sex differentiation.

Other genes associated with ovarian dysgenesis and POF include *LHX8*, *MCM8*, *MCM9*, *NOBOX*, and *FSHR* [46–48].

### 7.4.2 Exposure or overproduction of androgens

Increased androgen synthesis or exposure in utero leads to virilization of the female external genitalia in 46, XX individuals [25].

#### 7.4.2.1 Exposure to androgens of nonfetal origin

Maternal androgen-producing tumors including adrenal tumors and ovarian tumors can cause virilization of a female infant. There have been rare descriptions including a maternal luteoma of pregnancy, which caused virilization of both mother and child [53]. Drugs with androgenic activity including androgens, danazol, progestins, and potassium sparing diuretics can also virilization [15,53].

Placental aromatase deficiency is a rare cause of virilization in a female. Aromatase is responsible for converting androstenedione to estrone in the female ovaries. Placental aromatase protects the fetus from high circulating levels of androgens both from the fetal adrenal glands and maternal origin [54]. Placental aromatase deficiency is extremely rare presenting with maternal virilization in the third trimester and ambiguous genitalia in the affected females [54,55]. They have intact cortisol and aldosterone production so are not at risk for salt wasting. Mutations in *CYP19A1* on chromosome 15q21.2 are inherited in a recessive manner. There have been less than 20 cases described in literature.

#### 7.4.2.2 Steroid synthesis defects—overproduction of androgens

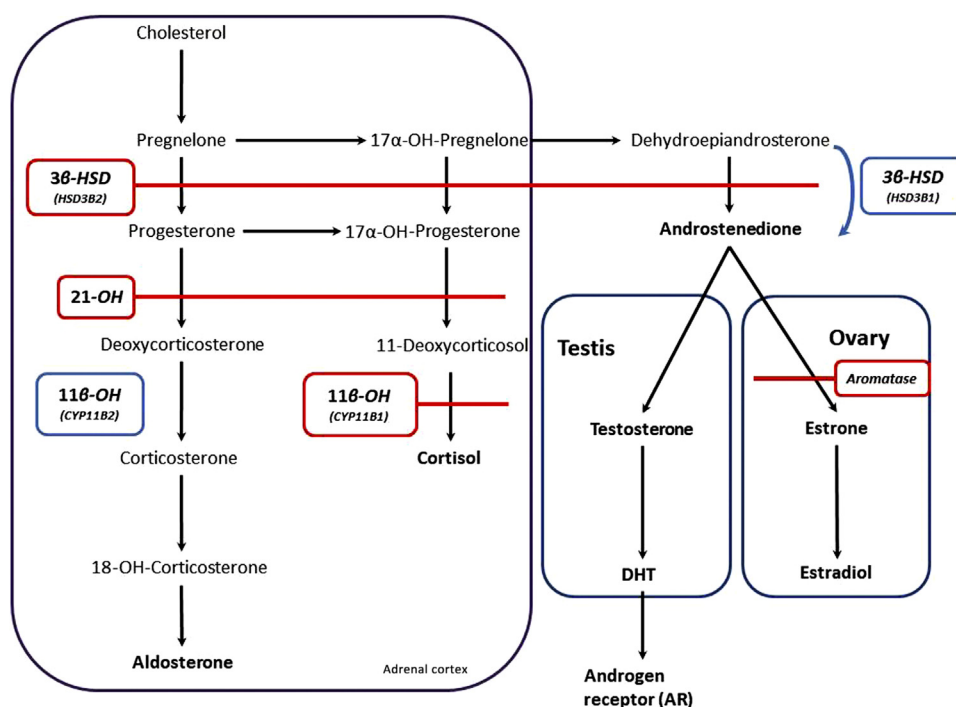
Steroid synthesis defects leading to CAH with androgen excess cause clitoral enlargement, partial or complete fusion of the labia majora and a short vagina [25]. The stages of virilization are staged according to the Prader score, stages 0–5 [55].

The most common cause of 46, XX disorders of sexual differentiation is 21-hydroxylase (21-OH) deficiency that occurs in 90% of cases [55] with a prevalence of 1:14,000–1:15,000 worldwide [54,55]. Mutations in the 21-OH gene (*CYP21A2*) on chromosome 6p21.1 are responsible for the condition and inherited in a recessive manner. The severity of the disease correlates generally to the degree of enzyme activity. Patients with classic 21(OH) deficiency have <5% activity and CAH that presents with in utero virilization of external genitalia in females and the risk of severe salt wasting in the neonatal period. Non-classic 21(OH) deficiency have >15% residual activity and present with androgen excess in adolescence and early adulthood [54].

In approximately 70% patients with classic 21(OH) deficiency present with salt wasting during the neonatal period [15]. Patients present with high 17 $\alpha$ -hydroxyprogesterone, androstenedione, and testosterone (Fig. 7.6). Infants with salt wasting present with low sodium, high potassium, and high renin at the end of the first week [15]. The increased androgens lead to virilization of the female external genitalia in 46, XX fetuses, and this occurs from an early gestational age [54]. The degree of virilization varies from mild clitoromegaly to complete masculinization [15].

Non-classic 21(OH) deficiency is more common than the classic form with an incidence of 1:300. In the Ashkenazi Jewish population the incidence is as high as 1:27 [54]. They generally present in adolescence with premature pubarche, acne, hirsutism, and irregular menses. The presentation is very similar to polycystic ovarian syndrome.

11 $\beta$ -OH is present in the adrenals and defects lead to CAH similar to the phenotype seen in the classical CAH. This is the second most common cause of CAH. Compared to 21-OH deficiency, it is quite rare with an incidence of <1:100,000 births. Mutations in the *CYP11B1* gene on chromosome 8q21–q22 are inherited in a recessive manner and lead to essentially all loss of enzyme activity [54]. Jews of Moroccan ancestry have a high prevalence due to a founder mutation, R448H, which is also the most common mutation in *CYP11B1* [54]. Patients with 11 $\beta$ -OH deficiency present with virilization of the external genitalia due androgen excess with increased 11-deoxycortisol and 17 $\alpha$ -hydroxypregnenolone (Fig. 7.5) [15]. However, they are at a decreased risk of salt wasting due to intact aldosterone production. Aldosterone synthesis is mediated by the *CYP11B2* gene, which is homologous to *CYP11B1* and located within 40 kb [54]. This leads to normal levels of deoxycorticosterone and aldosterone. There have been a wide range of blood pressure phenotypes described [25,54].



**FIGURE 7.6** Steroid hormone synthesis pathway in 46, XX DSD.

11 $\beta$ -OH, 11 $\beta$ -Hydroxylase; 21-OH, 21-hydroxylase; 3 $\beta$ -HSD, 3 $\beta$ -hydroxysteroid dehydrogenase; DSD, disorder of sex development. Source: Adapted from Wherrett, DK. Approach to the infant with a suspected disorder of sex development. *Pediatr Clin N Am* 2015;62(4):983–99.

Rarer forms of steroid synthesis defects that cause CAH include 3 $\beta$ -hydroxysteroid dehydrogenase (3 $\beta$ -HSD) deficiency and 11 $\beta$ -hydroxylase (11 $\beta$ -OH) deficiency. There are two enzymes that mediate 3 $\beta$ -HSD. Type 1 (*HSD3B1*) is present in the placenta, skin, liver, and peripheral tissues and is intact. Increased levels dehydroepiandrosterone are converted to androstenedione and along to testosterone and DHT through *HSD3B1* [54]. The degree of virilization is moderate compared to 21-OH deficiency presenting as mild to moderate clitoromegaly and rarely labiosacral fusion [54,55].

Type 2 (*HSD3B2*) deficiency is a rare condition occurring in 1:1,000,000 births. This condition presents with severe salt wasting similar to 21-OH deficiency, but the degree of virilization is less severe [54,55]. Mutations in the *HSD3B2* gene, located on chromosome 1p13, are inherited in a recessive manner with the severity of the disease correlating with the degree of enzyme activity [54,55]. Type 2 deficiency leads to impaired aldosterone and cortisol production with increased pregnenolone, 17(OH) pregnenolone, and increased dehydroepiandrosterone on investigation (Fig. 7.6) [15,55]. This leads to salt wasting, hyperkalemia, and volume depletion similar to 21(OH) deficiency. Virilization occurs in a similar fashion to 11 $\beta$ (OH) deficiency.

## 7.5 Investigations

Investigations should be done in all cases with genital abnormalities keeping in mind that in the most severe cases, which include 46, XY with complete feminization and 46, XX with complete masculinization, are difficult to diagnose as phenotypically abnormal. The investigative yield in isolated grade 1 (glandular) and 2 (penile shaft) hypospadias, especially when associated with intrauterine growth restriction (IUGR) is low [56]; however, since there is no objective way of differentiating IUGR–hypospadias syndrome from PAIS and mild 5 $\alpha$ -reductase deficiency, DNA analysis for these conditions are recommended. Investigations are also indicated when the hypospadias is associated with bilateral or unilateral undescended testes, males with micropenis, clitoromegaly in an apparently female infant with or without posterior fusion of the labia majora or an apparently male infant with isolated bilateral undescended testes.

The initial investigation should include a careful physical examination to determine if the DSD is isolated or nonisolated (Fig. 7.1). The results of maternal serum screening and prenatal ultrasound findings as well as the



birth weight are important in view of the association between IUGR and undervirilization in chromosomal male fetuses/newborns. Examination of the external genitalia should include assessment of the labioscrotal folds, phallic structure, urogenital openings, and position and volume of the gonads. Palpable gonads in the inguinal canal and labioscrotal folds are always testis although ovotestis cannot be ruled out. The labioscrotal folds should be assessed for pigmentation, rugation, asymmetry, and fusion. The phallic structure should be assessed for length, breadth, chordee, and relationship with the labioscrotal folds looking for complete or incomplete penoscrotal transposition. High insertion of the labioscrotal folds (above the penis) as well as “buried penis” should be differentiated from micropenis. The normal penile length at term is 3.5 cm with 2.5 cm being at  $-2SD$ . The normal clitoral length at term is 2–8.5 mm and breadth 2–6 mm [57]. The perineum should be examined for the number and position of the opening and compared to the Prader scale [58].

Investigations should be targeted to answer the following questions:

1. What is the chromosome sex?
2. What is the gonadal sex?
3. What are the findings on examination of the internal genitalia using ultrasound? Is there a uterus or gonads?
4. What are the findings on examination of the external genitalia?

Since gonads in the inguinal canal and labioscrotal folds are almost always, testis finding them in these positions usually indicates XY. However, the chromosome sex should be determined using quantitative fluorescent PCR or fluorescent in situ hybridization analysis and completed with microarray analysis looking for submicroscopic changes involving the SRY gene as well as in the other chromosomes [59].

The most common cause of ambiguous genitalia in female is CAH, and thus if the chromosome sex is XX, CAH should be ruled out by assessing the 17-hydroxyprogesterone (17-OH) and renin at 48 hours of age (after the surge of adrenal hormones at birth). The sodium and potassium blood levels will become abnormal only in the second week of life.

If the chromosome sex is XY, the testosterone, LH, and FSH levels should be measured to arrest the failure in the production of testosterone and DHT. Serum and DNA should always be banked in case further investigations will be needed. Obtaining information and answering the previous four questions will direct the secondary investigations that include assessment of adrenal function, testicular function, and internal genitalia using a genitogram. This is performed by injecting a radiopaque dye injected into the perineal opening [15].

Genetic testing plays an important role in finding the etiology and to provide counseling regarding recurrence risks and parental prenatal/preimplantation options in future pregnancies.

The use of a customized DSD genetic panel can interrogate multiple genes simultaneously and accelerate the diagnostic process. Using massive parallel sequencing with 64 known diagnostic DSD genes and candidate genes on a cohort of 278 patients with 46, XY DSD and 48 with 46, XX DSD of an unknown etiology, Eggers et al. found a likely genetic diagnosis in 43% of patients with 46, XY DSD and 17% of patients with 46, XX DSD [60]. If no gene mutation is being identified, whole exome/genome sequencing is recommended.

## 7.6 Gender assignment

In most newborns with DSD the findings on physical examination (external phenotypic sex) and the investigative results will enable a quick and easy decision regarding the sex of rearing. The inability to determine the gender role and gender identity in newborns has led to the suggestion by some to delay the surgical intervention to a time when the patient is able to make his/her decision. However, most experts in the field agree that a decision regarding the sex of rearing has to be made in infancy and surgical procedures to be carried out accordingly.

In cases with substantial ambiguity such as in PAIS, mixed gonadal dysgenesis, ovotesticular DSD, and females with CAH and severe masculinization, the sex of rearing can be more difficult and complicated. In these cases a multidisciplinary DSD team, which consists of an endocrinologist, urologist, geneticist, gynecologist, social worker, psychologist, and psychiatrist with an expertise in the field of DSD, should examine the baby, plan, and carry out the needed investigations. Afterward, they can interact and present the information to the parents to come a decision about sex of rearing. Factors that have to be taken into consideration include the necessary surgical treatment and the postsurgical external genital appearance, the internal genitalia, gonadal function and potential for fertility, the need for hormonal treatment to induce puberty and hormone replacement therapy and the wishes of the family. The team should use simple to understand terminology as well as visual aids to explain to the parents the findings and the options, so that they can make an informed decision.



It is important to keep in mind that the psychosocial sex (gender identity and gender role) is unknown at the time of initial presentation, and the decision regarding the sex of rearing is being made based on the findings on physical examination and the results of the laboratory investigation. The role of gender can vary depending on the condition. Information obtained from studies on girls prenatally exposed to high levels of androgens as in CAH show a more masculine gender role [61,62] and effect on sexual orientation [63,64]. This should be taken into account and shared with the parents as well as the findings on long term follow-up. Many cases with 46, XY DSD caused by *SRD5A2* gene mutations were raised as females decided to change to male gender at puberty. In cases with PAIS, studies have shown the difficulties in repairing severe hypospadias. This information should be shared with the parents when such a diagnoses is being made [65].

## 7.7 Gonadal cancer risk

The decision to pursue gonadectomy should be made based on the potential normal function of the gonad (hormonal and fertility potential), the decided sex of rearing as well as the risk for malignant degeneration. The presence of a Y chromosome in a dysgenetic gonad is associated with high risk for malignancy (gonadoblastoma, dysgerminoma, and germ cell neoplasia in situ [66,67]). With complete androgen insensitivity the removal of the gonad prior to puberty is controversial in view of the data indicating low risk for malignancy until early adult years [68,69]. Hence, some women with CAIS elect to keep gonads in situ. Laparoscopic gonadopexy to situate the gonads in a fixed position near the anterior abdominal wall to allow ultrasound surveillance and biopsies, if needed, may be an option for patients wishing to avoid gonadectomy [68,69]. Delayed surgery can help in involving the patient in the decision-making [67,70,71].

## 7.8 Conclusion

Normal sex development includes determination of chromosome sex, gonadal sex, development of internal and external genitalia, and psychosocial sex is a complex and well-orchestrated process, which involves many genetic and epigenetic components; many of them are yet to be discovered. DSD is an etiologically heterogeneous group of disorders but with a major continuous lifelong impact on patients and their families. To minimize potential harm the investigations should be done by a DSD team. The findings and current knowledge have to be presented to the parents, so they can make an informed decision. The patients and their families have to be followed and provided with thoughtful and respectful care by all involved, with all questions answered to the best of the DSD team knowledge. This should continue beyond childhood into puberty and adulthood to minimize the risk and maximize the benefit in order to achieve optimal psychosocial well-being, sexual satisfaction, and fertility in view of the patients' gender role and identity [72].

## 7.9 Resources Web-based resources

Web-based educational resources for families include: <[www.aboutkidshealth.ca/En/HowTheBodyWorks/SexDevelopmentAnOverview](http://www.aboutkidshealth.ca/En/HowTheBodyWorks/SexDevelopmentAnOverview)>. This website provides detailed graphically illustrated explanations of sex development and DSDs that health professionals can use when working with families.

### 7.9.1 Support groups

Androgen Insensitivity Syndrome: Differences of Sex Development Support Group. <<http://aisdsd.org/>>.

CARES Foundation: Congenital Adrenal Hyperplasia Research, Education and Support. <<http://www.cares-foundation.org>>.

Hypospadias and Epispadias Association. <<http://heainfo.org>>.

## References

- [1] Hughes IA, Nihoul-Fekete C, Thomas B, Cohen-Kettenis PT. Consequences of the ESPE/LWPES guidelines for diagnosis and treatment of disorders of sex development. *Best Pract Res Clin Endocrinol Metab* 2007;21(3):351–65.
- [2] Hughes IA, Houk C, Ahmed SF, Lee PA. Consensus statement on management of intersex disorders. *J Pediatr Urol* 2006;2(3):148–62.

- [3] Yao HH. The pathway to femaleness: current knowledge on embryonic development of the ovary. *Mol Cell Endocrinol* 2005;230(1–2):87–93.
- [4] Zhao F, Franco HL, Rodriguez KF, et al. Elimination of the male reproductive tract in the female embryo is promoted by COUP-TFII in mice. *Science (New York, NY)* 2017;357(6352):717–20.
- [5] Childs AJ, Cowan G, Kinnell HL, Anderson RA, Saunders PT. Retinoic acid signalling and the control of meiotic entry in the human fetal gonad. *PLoS One* 2011;6(6):e20249.
- [6] Nielsen J, Wohler M. Chromosome abnormalities found among 34,910 newborn children: results from a 13-year incidence study in Arhus, Denmark. *Hum Genet* 1991;87(1):81–3.
- [7] Cameron-Pimblett A, La Rosa C, King TFJ, Davies MC, Conway GS. The Turner syndrome life course project: karyotype-phenotype analyses across the lifespan. *Clin Endocrinol* 2017;87(5):532–8.
- [8] Gravholt CH, Andersen NH, Conway GS, et al. Clinical practice guidelines for the care of girls and women with Turner syndrome: proceedings from the 2016 Cincinnati International Turner Syndrome Meeting. *Eur J Endocrinol* 2017;177(3):G1–g70.
- [9] Schoemaker MJ, Swerdlow AJ, Higgins CD, Wright AF, Jacobs PA. Cancer incidence in women with Turner syndrome in Great Britain: a national cohort study. *Lancet Oncol* 2008;9(3):239–46.
- [10] Flannigan R, Patel P, Paduch DA. Klinefelter syndrome. The effects of early androgen therapy on competence and behavioral phenotype. *Sex Med Rev* 2018;6(4):595–606.
- [11] Telvi L, Lebbat A, Del Pino O, Barbet JP, Chaussain JL. 45, X/46, XY mosaicism: report of 27 cases. *Pediatrics* 1999;104(2 Pt 1):304–8.
- [12] Tho SP, Jackson R, Kulharya AS, Reindollar RH, Layman LC, McDonough PG. Long-term follow-up and analysis of monozygotic twins concordant for 45, X/46, XY peripheral blood karyotype but discordant for phenotypic sex. *Am J Med Genet, A* 2007;143a(21):2616–22.
- [13] Tosson H, Rose SR, Gartner LA. Description of children with 45, X/46, XY karyotype. *Eur J Pediatr* 2012;171(3):521–9.
- [14] Lindhardt Johansen M, Hagen CP, Rajpert-De Meyts E, et al. 45, X/46, XY mosaicism: phenotypic characteristics, growth, and reproductive function—a retrospective longitudinal study. *J Clin Endocrinol Metab* 2012;97(8):E1540–9.
- [15] Wherrett DK. Approach to the infant with a suspected disorder of sex development. *Pediatr Clin N Am* 2015;62(4):983–99.
- [16] Rey R, Josso N, Racine C. Sexual differentiation. In: De Groot LJ, Chrousos G, Dungan K, et al., editors. *Endotext*. South Dartmouth, MA: MDText.com, Inc; 2000.
- [17] Berta P, Hawkins JR, Sinclair AH, et al. Genetic evidence equating SRY and the testis-determining factor. *Nature* 1990;348(6300):448–50.
- [18] Sekido R, Lovell-Badge R. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* 2008;453(7197):930–4.
- [19] Benko S, Gordon CT, Mallet D, et al. Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. *J Med Genet* 2011;48(12):825–30.
- [20] Vetro A, Ciccone R, Giorda R, et al. XX males SRY negative: a confirmed cause of infertility. *J Med Genet* 2011;48(10):710–12.
- [21] Bagheri-Fam S, Bird AD, Zhao L, et al. Testis determination requires a specific FGFR2 isoform to repress FOXL2. *Endocrinology* 2017;158(11):3832–43.
- [22] Song Y, Fan L, Gong C. Phenotype and molecular characterizations of 30 children from china with NR5A1 mutations. *Front Pharmacol* 2018;9:1224.
- [23] Tantawy S, Mazen I, Soliman H, et al. Analysis of the gene coding for steroidogenic factor 1 (SF1, NR5A1) in a cohort of 50 Egyptian patients with 46, XY disorders of sex development. *Eur J Endocrinol* 2014;170(5):759–67.
- [24] Yagi H, Takagi M, Kon M, Igarashi M, Fukami M, Hasegawa Y. Fertility preservation in a family with a novel NR5A1 mutation. *Endocr J* 2015;62(3):289–95.
- [25] Barbaro M, Wedell A, Nordenstrom A. Disorders of sex development. *Semin Fetal Neonatal Med* 2011;16(2):119–27.
- [26] Firth HV, Hurst JA. Oxford desk reference. New York: Oxford University Press; 2017.
- [27] Biason-Lauber A, Konrad D, Meyer M, DeBeaufort C, Schoenle EJ. Ovaries and female phenotype in a girl with 46, XY karyotype and mutations in the CBX2 gene. *Am J Hum Genet* 2009;84(5):658–63.
- [28] Loke J, Pearlman A, Radi O, et al. Mutations in MAP3K1 tilt the balance from SOX9/FGF9 to WNT/beta-catenin signaling. *Hum Mol Genet* 2014;23(4):1073–83.
- [29] Sugie K, Futamura N, Suzumura A, Tate G, Umehara F. Hereditary motor and sensory neuropathy with minifascicle formation in a patient with 46XY pure gonadal dysgenesis: a new clinical entity. *Ann Neurol* 2002;51(3):385–8.
- [30] Ahmed SF, Bashamboo A, Lucas-Herald A, McElreavey K. Understanding the genetic aetiology in patients with XY DSD. *Br Med Bull* 2013;106:67–89.
- [31] Yatsenko SA, Witchel SF. Genetic approach to ambiguous genitalia and disorders of sex development: what clinicians need to know. *Semin Perinatol* 2017;41(4):232–43.
- [32] HersHKovitz E, Parvari R, Wudy SA, et al. Homozygous mutation G539R in the gene for P450 oxidoreductase in a family previously diagnosed as having 17,20-lyase deficiency. *J Clin Endocrinol Metab* 2008;93(9):3584–8.
- [33] Yang Z, Ye L, Wang W, et al. 17beta-Hydroxysteroid dehydrogenase 3 deficiency: three case reports and a systematic review. *J Steroid Biochem Mol Biol* 2017;174:141–5.
- [34] Massanyi EZ, Dicarlo HN, Migeon CJ, Gearhart JP. Review and management of 46, XY disorders of sex development. *J Pediatr Urol* 2013;9(3):368–79.
- [35] Mongan NP, Tadokoro-Cuccaro R, Bunch T, Hughes IA. Androgen insensitivity syndrome. *Best Pract Res Clin Endocrinol Metab* 2015;29(4):569–80.
- [36] Boehmer AL, Brinkmann O, Bruggenwirth H, et al. Genotype versus phenotype in families with androgen insensitivity syndrome. *J Clin Endocrinol Metab* 2001;86(9):4151–60.
- [37] Hornig NC, de Beaufort C, Denzer F, et al. A recurrent germline mutation in the 5'UTR of the androgen receptor causes complete androgen insensitivity by activating aberrant uORF translation. *PLoS One* 2016;11(4):e0154158.
- [38] Kansakoski J, Jaaskelainen J, Jaaskelainen T, et al. Complete androgen insensitivity syndrome caused by a deep intronic pseudoexon-activating mutation in the androgen receptor gene. *Sci Rep* 2016;6:32819.

- [39] Hornig NC, Ukat M, Schweikert HU, et al. Identification of an AR mutation-negative class of androgen insensitivity by determining endogenous AR activity. *J Clin Endocrinol Metab* 2016;101(11):4468–77.
- [40] He X, Semenov M, Tamai K, Zeng X. LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way. *Development (Cambridge, England)* 2004;131(8):1663–77.
- [41] Nusse R. Wnts and Hedgehogs: lipid-modified proteins and similarities in signaling mechanisms at the cell surface. *Development (Cambridge, England)* 2003;130(22):5297–305.
- [42] Jordan BK, Mohammed M, Ching ST, et al. Up-regulation of WNT-4 signaling and dosage-sensitive sex reversal in humans. *Am J Hum Genet* 2001;68(5):1102–9.
- [43] McLaren A. Meiosis and differentiation of mouse germ cells. *Symp Soc Exp Biol* 1984;38:7–23.
- [44] Mandel H, Shemer R, Borochowitz ZU, et al. SERKAL syndrome: an autosomal-recessive disorder caused by a loss-of-function mutation in WNT4. *Am J Hum Genet* 2008;82(1):39–47.
- [45] Biason-Lauber A, Konrad D. WNT4 and sex development. *Sex Dev: Genet Mol Biol Evol Endocrinol Embryol Pathol Sex Determination Differ* 2008;2(4–5):210–18.
- [46] Desai S, Wood-Trageser M, Matic J, et al. MCM8 and MCM9 nucleotide variants in women with primary ovarian insufficiency. *J Clin Endocrinol Metab* 2017;102(2):576–82.
- [47] Katari S, Wood-Trageser MA, Jiang H, et al. Novel inactivating mutation of the FSH receptor in two siblings of Indian origin with premature ovarian failure. *J Clin Endocrinol Metab* 2015;100(6):2154–7.
- [48] Ren Y, Suzuki H, Jagarlamudi K, et al. Lhx8 regulates primordial follicle activation and postnatal folliculogenesis. *BMC Biol* 2015;13:39.
- [49] Uhlenhaut NH, Jakob S, Anlag K, et al. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* 2009;139(6):1130–42.
- [50] Ellsworth BS, Egashira N, Haller JL, et al. FOXL2 in the pituitary: molecular, genetic, and developmental analysis. *Mol Endocrinol (Baltimore, MD)* 2006;20(11):2796–805.
- [51] Pannetier M, Chassot AA, Chaboissier MC, Pailhoux E. Involvement of FOXL2 and RSPO1 in ovarian determination, development, and maintenance in mammals. *Sex Dev: Genet Mol Biol Evol Endocrinol Embryol Pathol Sex Determination Differ* 2016;10(4):167–84.
- [52] Tomaselli S, Megiorni F, Lin L, et al. Human RSPO1/R-spondin1 is expressed during early ovary development and augments beta-catenin signaling. *PLoS One* 2011;6(1):e16366.
- [53] Spitzer RF, Wherrett D, Chitayat D, et al. Maternal luteoma of pregnancy presenting with virilization of the female infant. *J Obstet Gynaecol Can* 2007;29(10):835–40.
- [54] Auchus RJ, Chang AY. 46, XX DSD: the masculinised female. *Best Pract Res Clin Endocrinol Metab* 2010;24(2):219–42.
- [55] Paris F, Gaspari L, Philibert P, Maimoun L, Kalfa N, Sultan C. Disorders of sex development: neonatal diagnosis and management. *Endocr Dev* 2012;22:56–71.
- [56] Yinon Y, Kingdom JC, Proctor LK, et al. Hypospadias in males with intrauterine growth restriction due to placental insufficiency: the placental role in the embryogenesis of male external genitalia. *Am J Med Genet, A* 2010;152a(1):75–83.
- [57] Oberfield SE, Mondok A, Shahrivar F, Klein JF, Levine LS. Clitoral size in full-term infants. *Am J Perinatol* 1989;6(4):453–4.
- [58] Allen L. Disorders of sexual development. *Obstet Gynecol Clin N Am* 2009;36(1):25–45.
- [59] Tannour-Louet M, Han S, Corbett ST, et al. Identification of de novo copy number variants associated with human disorders of sexual development. *PLoS One* 2010;5(10):e15392.
- [60] Eggers S, Sadedin S, van den Bergen JA, et al. Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. *Genome Biol* 2016;17(1):243.
- [61] Mathews GA, Fane BA, Conway GS, Brook CG, Hines M. Personality and congenital adrenal hyperplasia: possible effects of prenatal androgen exposure. *Horm Behav* 2009;55(2):285–91.
- [62] Nordenstrom A, Servin A, Bohlin G, Larsson A, Wedell A. Sex-typed toy play behavior correlates with the degree of prenatal androgen exposure assessed by CYP21 genotype in girls with congenital adrenal hyperplasia. *J Clin Endocrinol Metab* 2002;87(11):5119–24.
- [63] Frisen L, Nordenstrom A, Falhammar H, et al. Gender role behavior, sexuality, and psychosocial adaptation in women with congenital adrenal hyperplasia due to CYP21A2 deficiency. *J Clin Endocrinol Metab* 2009;94(9):3432–9.
- [64] Meyer-Bahlburg HF, Dolezal C, Baker SW, New MI. Sexual orientation in women with classical or non-classical congenital adrenal hyperplasia as a function of degree of prenatal androgen excess. *Arch Sex Behav* 2008;37(1):85–99.
- [65] Lucas-Herald A, Bertelloni S, Juul A, et al. The long-term outcome of boys with partial androgen insensitivity syndrome and a mutation in the androgen receptor gene. *J Clin Endocrinol Metab* 2016;101(11):3959–67.
- [66] Cools M, Looijenga L. Update on the pathophysiology and risk factors for the development of malignant testicular germ cell tumors in complete androgen insensitivity syndrome. *Sex Dev: Genet Mol Biol Evol Endocrinol Embryol Pathol Sex Determination Differ* 2017;11(4):175–81.
- [67] Wunsch L, Holterhus PM, Wessel L, Hiort O. Patients with disorders of sex development (DSD) at risk of gonadal tumour development: management based on laparoscopic biopsy and molecular diagnosis. *BJU Int* 2012;110(11 Pt C):E958–65.
- [68] Chaudhry S, Tadokoro-Cuccaro R, Hannema SE, Acerini CL, Hughes IA. Frequency of gonadal tumours in complete androgen insensitivity syndrome (CAIS): a retrospective case-series analysis. *J Pediatr Urol* 2017;13(5):498.e491–6.
- [69] Dohnert U, Wunsch L, Hiort O. Gonadectomy in complete androgen insensitivity syndrome: why and when? *Sex Dev: Genet Mol Biol Evol Endocrinol Embryol Pathol Sex Determination Differ* 2017;11(4):171–4.
- [70] Mouriquand PD, Gorduza DB, Gay CL, et al. Surgery in disorders of sex development (DSD) with a gender issue: if (why), when, and how? *J Pediatr Urol* 2016;12(3):139–49.
- [71] Patel V, Casey RK, Gomez-Lobo V. Timing of gonadectomy in patients with complete androgen insensitivity syndrome-current recommendations and future directions. *J Pediatr Adolesc Gynecol* 2016;29(4):320–5.
- [72] Raveenthiran V. Neonatal sex assignment in disorders of sex development: a philosophical introspection. *J Neonatal Surg* 2017;6(3):58.



# Molecular systems in cardiovascular developmental disorders

Alan Graham Stuart

Honorary Associate Professor of Sports and exercise Cardiology, University of Bristol, Bristol, United Kingdom;  
Consultant Cardiologist, Bristol Heart Institute, Bristol, United Kingdom

## 8.1 Introduction

The term *cardiovascular developmental disorder* can be defined as “any structural abnormality of the cardiovascular system, which is a consequence of abnormal development from early fetal life to adulthood.” The study of cardiovascular developmental disorders has been a major focus of laboratory, clinical, and epidemiological investigation for centuries, and advances in our understanding of the pathogenesis and consequences of cardiovascular developmental disorders have led to outstanding improvements in prognosis for patients with congenital heart disease (CHD). Thus 25 years ago, a child diagnosed with the severe congenital heart defect hypoplastic left heart syndrome (HLHS) in utero would usually die within 1 week of extrauterine life. In 2018 a child born with HLHS who receives modern medical and surgical care has an excellent chance ( $> 80\%$ ) of reaching adulthood [1]. In parallel with the advances in medical and surgical care, there have been rapid improvements in our understanding of the molecular mechanisms associated with the development of congenital cardiovascular developmental disorders. It is the purpose of this chapter to examine these developments.

## 8.2 Historical overview

The earliest description of a congenital heart defect was by the Danish catholic bishop and anatomist Nicholas Stensen (1636–86) who noted a defect in the ventricular septum and narrowed pulmonary valve during the post-mortem examination of a fetus [2]. Over the years it became clear that congenital cardiovascular malformations are common and occur in around 1% of live births with severe, life-threatening defects occurring in 0.2%–0.3% live births [3,4]. By the end of the 20th century the introduction of safe and effective cardiac surgery had transformed the outlook of many patients with CHD. The first cardiac surgical repair, ligation of a patient’s arterial duct, was performed in 1938 by Goss [5]. However, it was the development of the ability to perfuse the heart and lungs while operating on the heart in the 1950s that led to the remarkable improvements in outcome. This technique allowed direct vision of the defect enabling even complex defects such as tetralogy of Fallot (TOF) to undergo surgical repair [6]. Sequential improvements in cardiopulmonary bypass technology, surgical technique, and postoperative care have resulted in a current 30-day postsurgical mortality of below 3% [7].

## 8.3 Normal development of the heart

A full description of the development of the cardiovascular system is beyond the scope of this chapter, and excellent descriptions can be found elsewhere [8,9]. In vertebrates this involves a series of molecular and

Days postconception	Cardiac development milestone
17–21	Formation of the cardiac crescent, cardiac jelly forms cardiac cavity, and endocardial plexus appears
22–32	Heart starts beating, cardiac looping starts, atria and ventricles balloon, septum primum/muscular ventricular septae start to form, cells appear in cardiac jelly, and AV cushions form
32–42	AV cushions oppose, pulmonary vein attaches, outflow tract ridges start to septate, primum septum closes, and epicardium forms
42–52	Secundum septum and sinus node appear, AV valves start to form and papillary muscles appear, and mural AV valve leaflets released
52–60	Coronary branches appear, cordae tendinae form, and septal leaflet of the tricuspid valve delaminates

**FIGURE 8.1** Approximate timeline of cardiac development in the human fetus.

morphological events beginning with migration of precardiac cells and assembly of paired cardiac crescents. This is followed by coalescence of the cardiac crescents to form a primitive cardiac tube, which in turn loops to ensure correct alignment of the cardiac chambers. The chambers then form with associated chamber septation, and finally the cardiac conduction system and coronary artery vessels develop. Many of the recent advances in our understanding of cardiac development have been achieved using molecular genetic techniques. Thus the development of techniques enabling an irreversible label to be attached to cells, which express a specific gene, has allowed the development of fate maps that can be used to demonstrate the cellular origin of different components of the heart. In turn, this can be used to inform our understanding of the development of cardiovascular malformations.

The main components of human cardiovascular development are laid down within the first 8 weeks of fetal life. The transformation of the primitive cardiac tube to a four-chambered pump, which works in series with low pressure—recipient chambers (atria) filling high pressure—pumping chambers (ventricles), requires a precise balance between cellular growth, cell differentiation, and apoptosis. Any disruption to a critical event, such as looping, septation, and chamber formation, can lead to a major abnormality of the completed cardiac structure and function. An approximate timeline of critical events is shown in [Fig. 8.1](#).

## 8.4 Advances in technology

Blue has estimated that there are likely to be at least 400 genes involved in the etiology of CHD with many yet to be discovered [\[10\]](#). Much of our understanding of genetics in CHD has arisen from an evaluation of familial CHD. Using techniques such as linkage analysis and candidate gene approaches, investigators were able to identify genes such as NKX2.5, TBX5, and NOTCH1 associated with CHD [\[11–13\]](#). In turn, these findings led to further work that evaluated the underlying molecular mechanisms. For example, using a zebrafish model, Sultana demonstrated that NKX2.5 is a downstream mediator of connexin Ecx-mediated signaling and thus promotes parallel alignment of myofibrils and subsequent correct heart morphogenesis [\[14\]](#). However, this approach requires large families with many affected individuals, which is uncommon in CHD.

### 8.4.1 Genome-wide association studies

In this method, genome-wide search is carried out for small genetic variations (single-nucleotide polymorphisms—SNPs), which are more common in individuals with a particular disease [\[15\]](#). These studies have been helpful in evaluating the genetic influences in complex traits as opposed to single-gene disorders. Genome-wide association studies (GWAS) studies have been used increasingly in CHD, for example, in TOF, conotruncal defects, and septal defects [\[16–18\]](#). However, the results need to be interpreted with caution. Thus preliminary GWAS studies associated the MTHFR C677T polymorphism (which influences plasma folate levels) with the presence of CHD. This link supported the hypothesis that abnormalities of folate metabolism contribute to the molecular changes resulting in CHD. However, a more robust study and metanalysis of the literature demonstrated that the MTHFR C677T polymorphism does not link to CHD and this was in part due to publication bias [\[19\]](#).



### 8.4.2 Whole genome and exome sequencing

The advent of next-generation sequencing (NGS) technology, such as whole-exome sequencing (WES) and whole-genome sequencing (WGS), is a major advance and allows an assessment to be made of multiple sources of genetic variation including common and rare variants, deletions and duplications, and copy-number variations (CNVs). WGS attempts to sequence the entire genome, whereas WES focuses only on protein-coding genes (around 1.5% of the genome). There are many relative merits of each approach. These are described in detail elsewhere [20]. However, this technology enables the study of relatively small numbers of individuals in families with a clear phenotype, as opposed to the much larger populations required for GWAS. Depending on the diagnostic question, some laboratories will restrict the exons being evaluated to a panel of genes known to be associated with the suspected disorder.

Pawlak reviewed the use of NGS in cardiac development [21]. WES has enabled the identification of novel mutations associated with many cardiac developmental anomalies including atrial septal defects (ASDs), bicuspid aortic valve (BAV), patent ductus arteriosus (PDA), TOF, left-ventricular outflow tract obstructions, and left-ventricular noncompaction. Targeted NGS studies provide a more cost-effective strategy. Using a pooling and targeted exon-resequencing strategy, Bonachea identified 97 candidate genes in BAV patients (including 33 with *in silico* prediction to be disease-causing mutations) within GATA4/5, APC, JAG1, NOTCH1/2/3, PAX8, SOX9, TBX5, and WNT4—of these, 26 genes were not previously associated with BAV [22]. This targeted strategy confirmed the complex polygenic nature of BAV and was estimated to have saved \$39,350 compared to an unpooled strategy [22].

Although novel mutations have usually been found within genes already known to be associated with CHD, the NGS approach has enabled a new insight into many molecular mechanisms including the role of epigenetic factors and noncoding RNA. This is discussed in more detail later. Using NGS, often in combination with animal models, further insight has also been gained into precise molecular mechanisms underlying the development of CHD. For example, novel genes have been identified in HLHS by combining mouse forward genetics and WES. These findings, confirmed in functional studies in the mouse and zebrafish, demonstrated that *Sap10* mediates left-ventricular hypoplasia, whereas *Pcdha9* increases the penetrance of aortic valve malformations [23].

## 8.5 Chromosomal aneuploidy and structural heart disease

A significant proportion of CHD is associated with numerical chromosome abnormalities (aneuploidy) (Table 8.1). It is a common sight in most congenital heart units to see one or two patients with Down syndrome under investigation or recovering from cardiac surgery. CHD is present in 44% of children with Down syndrome, and there is evidence that the prevalence is increasing as a consequence of improved medical management and progress in socioeconomic support for families [24,25]. The majority of Down syndrome patients have trisomy of chromosome 21, but this can be partial (mosaicism/translocation) in up to 6%. The molecular mechanism that causes CHD in Down syndrome is still not entirely clear, but recent work implicates a candidate region on 21q22.2 (0.96 Mb). This region is shared by most Down syndrome patients with CHD and contains three known protein-coding genes (*DSCAM*, *BACE2*, and *PLAC4*) and four known noncoding RNAs (*DSCAM-AS1*, *DSCAM-IT1*, *LINC00323*, and *MIR3197*) [26].

In a 26-year study of 346,831 consecutive pregnancies, Stoll reported 4005 cases with CHD of whom 354 (8.8%) had chromosomal abnormalities including 218 trisomies 21, and 99 had (2.5%) nonchromosomal-dysmorphic conditions [27]. Table 8.1 lists the common chromosomal anomalies linked to CHD.

### 8.5.1 Congenital heart disease and copy-number variations

CNVs are structural gene variants, which involve alterations in the number of copies of a specific region of DNA. These can be either deletions or duplications. CNVs range in size from single genes to millions of base pairs. Deletions tend to be more problematic than duplications because of the sensitivity to gene dosage in many genes that do not tolerate haploinsufficiency. Array comparative genome hybridization is a quick and efficient way of identifying CNVs in a test genome relative to a reference genome although NGS technologies can also be used.

Patients with CHD who have CNVs have a higher frequency of noncardiac anomalies and hence tend to have a poorer outlook. In addition to CHD, there are often abnormalities in intellectual development and growth.

**TABLE 8.1** Common chromosomal syndromes associated with congenital heart disease (aneuploidy and microdeletions).

Syndrome	Genetic anomaly	Cardiac anomaly	Other features	% of CHDs	Comments and genes implicated
Patau syndrome	Trisomy 18	ASD, VSD, HLHS, PDA	Multiple anomalies including severe developmental delay, polydactyly, and clefts	>80	Usually fatal in infancy
Edwards syndrome	Trisomy 13	ASD, VSD, HLHS, PDA, ToF, CoA, act	Multiple anomalies including severe developmental delay, biliary atresia, and omphalocele	>90	1:6000 live births. 90% die in infancy
Down syndrome	Trisomy 21	ASD, PDA, ToF, AS, AVSD, CoA	Variable developmental delay, hypothyroidism, brachycephaly, short stature	44	CHD is a major determinant of long-term outcome. Pulmonary hypertension is particularly aggressive when present. May have translocation or mosaicism (see text)
Klinefelter syndrome	47XXY	ASD, PDA, MVP	Variable developmental delay, tall stature, hyogonadism	50	Tendency toward diastolic dysfunction, increased intima–media thickness, increased thromboembolic events, and chronotropic incompetence with age. These findings are not reversed by testosterone therapy
Turner syndrome	45 XO	BAV, CoA, HLHS, dil Ao	Short stature, lymphedema in infancy, shield chest, reduced fertility	35	Progressive ascending aortopathy with risk of dissection mandates long-term follow-up. Haploinsufficiency of short arm of X chromosome implicated in cardiac phenotype
William syndrome	7q11.23 deletion	Suprav AS, Branch PAS, BAV, CA stenoses	Variable developmental delay, stellate iris, “cocktail party personality” as child, infantile hypercalcemia, elfin facies	50–85	Branch PAS tends to improve and supravalar AS tends to increase in severity with age
DiGeorge syndrome/ Chrome 22 microdeletion syndrome	22q11.2	Conotruncal anomalies AS, BAV, IAA, CoA, SAT, ToF	Infantile hypocalcemia, T-cell deficiency, behavior/psychiatric disturbance as adults, renal anomalies, palate anomalies	80	Typical facial appearance in infancy. T-cell deficiency requires use of irradiated blood at surgery. Psychiatric disturbance may progress to psychosis. 40 genes present in commonly deleted region (including TBX1) but candidate gene studies inconclusive

AS, Aortic stenosis; ASD, atrial septal defect; AVSD, atrioventricular septal defect; BAV, bicuspid aortic valve; Branch PAS, branch pulmonary artery stenosis; CA stenosis, coronary artery stenoses; CoA, coarctation of aorta; dil Ao, ascending aortopathy; HLHS, hypoplastic left-heart syndrome; IAA, interrupted aortic arch; MVP, mitral valve prolapse; PDA, patent ductus arteriosus; SAT, single arterial trunk; Suprav AS, supravalar aortic stenosis, ToF, tetralogy of Fallot; VSD, ventricular septal defect.

The most common deletion syndrome and one of the most common CNV-associated CHD syndromes is chromosome 22q11 microdeletion syndrome. This has a prevalence of 1:6000 live births and is associated with conotruncal anomalies as well as many noncardiac features including renal anomalies, developmental delay, behavior disturbance ranging from an anxiety disorder to psychosis, and DiGeorge syndrome with hypocalcemia and T-cell immunodeficiency [28,29]. There are at least 40 genes in the commonly deleted region of Chr22q11 and although the precise molecular mechanism that results in CHD remains unclear, it appears that chromatin modification and gene–gene interaction may be important [30]. Xing has shown that the noncoding regulatory genome may also play a role with 18% of their group of CNV in association with CHD patients harboring potentially important microRNAs (miRNAs) [31].

## 8.6 Single-gene (Mendelian) disorders

Examples of some of the more common single-gene disorders associated with cardiovascular developmental disorders are described in Table 8.2.

**TABLE 8.2** Examples of single-gene defects associated with developmental cardiovascular conditions.

Disorder type	Cardiac anomaly	Other features	% of CHDs	Comments and genes implicated
RAS/MAPK syndromes	PS, AVSD CoA, ToF ASD, PDA VSD MV anomalies HCM	Lentigines Dysmorphism Bleeding diathesis Lymphedema Increased cancer risk	Up to 90% (Noonan syndrome)	Noonan syndrome Leopard syndrome Costello syndrome Cardiofaciocutaneous syndrome Genes involved PTPN11, SOS1, RAF1, KRAS, NRAS, RIT1, SHOC2, SOS2, BRAF
Transcription factor–related heart disease	ASD, VSD, ToF, TGA, AVSD, DORV, arrhythmias	Limb abnormalities	Up to 75% (Holt–Oram syndrome)	Holt–Oram syndrome TBX5, NKX2.5, GATA4, NKX2.6, ZIC3, TBX20, CRELD1, NOTCH1, MYH6, CITED2
CHARGE association	AVSD, conotruncal defects, aortic arch anomalies	Coloboma, ear anomalies, choanal atresia, cranial nerve defects, semicircular canal defects, mental retardation, hypogonadism	74%	75% CHD7 loss of function. Heart defects more frequent in truncating mutations. 20% have normal intelligence. External ear, semicircular canal defects, and cranial nerve dysfunction almost always present
Miscellaneous single-gene causes	ASD, TGA VSD, DORV CoA AVSD HRV, HLHS	Isolated CHD only	Variable	NKX2.5 GATA4 TBX20 NOTCH1 ELN CRELD1

ASD, Atrial septal defect; AVSD, atrioventricular septal defect; BAV, bicuspid aortic valve; CoA, coarctation of aorta; *dil Ao*, ascending aortopathy; DORV, double outlet right ventricle; HCM, hypertrophic cardiomyopathy; HLHS, hypoplastic left heart syndrome; HRV, hypoplastic right ventricle; MV anomalies, mitral valve anomalies; PDA, patent arterial duct; PS, pulmonary stenosis; *suprav AS*, supraaortic stenosis; TGA, transposition of great arteries; ToF, tetralogy of Fallot; VSD, ventricular septal defect.

### 8.6.1 RASopathies

The RASopathies, of which the most common is Noonan syndrome, are autosomal dominant and pleiotropic. They are characterized by mutations that act on signal transduction via the RAS–MAPK pathway. For example, PTPN11 mutations, which cause approximately 50% of Noonan syndrome, encode the protein tyrosine phosphatase (SHP2), which results in a gain of function with increased and prolonged phosphatase activity [32,33]. The SHP2 mutants prolong signal flux through the RAS/mitogen-activated protein kinase (ERK2/MAPK1) pathway in a ligand-dependent manner, which results in increased cell proliferation and the clinical phenotype [33].

### 8.6.2 Transcription factor–related disorders [34]

Holt–Oram syndrome is an uncommon syndrome affecting heart and limb development. It is the most common of the transcription factor–related disorders and has a prevalence of 1:100,000 [12]. Although rare, it is associated with a high incidence of CHD and has been used as a model to improve our understanding of the molecular mechanism associated with CHD—particularly through its association with TBX5. Fan et al. have demonstrated that TBX5 plays a role in pre-mRNA splicing with the protein SC35 identified as the affected splicing factor [35]. The most severe TBX mutation, G80R, with complete penetrance of the cardiac phenotype strongly affects pre-mRNA slicing although this is not seen in all TBX5 mutations. Thus TBX5 can play a role in transcriptional activation and pre-mRNA splicing. TBX5 can also affect chromatin remodeling, and different isoforms of TBX5 have different transcriptional properties such that TBX5c antagonizes TBX5a-activation of proproliferative signals such as IGF-1, FGF-10, and BMP4 [36,37]. These results provide new insight into Tbx5 regulation and function and emphasize the complex molecular interactions that occur to produce the Holt–Oram phenotype.

#### 8.6.2.1 CHARGE association

The CHARGE association was first described in 1981 and consists of multiple congenital malformations including a variable combination of Coloboma, Heart anomalies, Atresia choanae, Growth retardation, and

Ear abnormalities [38]. It is associated with a mutation in the CHD7 gene on chromosome 8q12, which, like TBX5, plays a role in gene transcription by affecting chromatin remodeling [39]. The precise functions of CHD7 have not been fully evaluated, but it seems to play a role in embryonic heart development. It has been demonstrated to regulate both nucleoplasmic and nucleolar genes by enhancer-mediated transcription and nucleolar RNA biogenesis. Moreover, Zentner demonstrated in a mouse model that both haploinsufficiency and complete loss of CHD7 lead to increased DNA methylation of the rRNA promoter and decreased rRNA expression [40]. CHD7 is also expressed in neural and pharyngeal arches, which may explain the noncardiac features.

### 8.6.3 Isolated congenital heart disease caused by single gene

The cause of most isolated CHD (80%) is probably polygenic or multifactorial. However, around 20% of isolated CHDs can be attributed to single-gene disorders. Thus most ASDs are not known to have a genetic origin, but a mutation in NKX2.5 explains 1%–4% of sporadic ASD and up to 12% of familial ASD [41]. This is an important association to recognize as not only are NKX2.5-associated ASD patients at risk of progressive conduction disease, but also a high risk of sudden cardiac death due to ventricular arrhythmias and a defibrillator may be indicated [42]. Moreover, even in the absence of an ASD, an NKX2.5 mutation can cause progressive conduction disorder. Mutation in the zinc finger transcription factor GATA4 gene is also associated with ASD but in the absence of conduction defects [43]. GATA4 mutations have also associated with isolated ventricular septal defect (VSD), pulmonary stenosis, atrioventricular (AV) septal defect, and dilated cardiomyopathy. The study of genes, such as TBX5, GATA4, and NKX2.5, has led to the recognition of certain phenotypic patterns such that structural features of these genes have been proposed as predictors of the CHD clinical subtype [44]. Thus CHD variants of NKX2.5 are located primarily in the tinman and homeodomain with the AV block/ASD combination correlating with the AV–block clinical subtype. Similarly, the VSD-associated GATA4 variant tends to be associated with a terminal end mutation [44].

The systematic study of families with nonsyndromic, highly penetrant, and autosomal-dominant CHD has enabled the identification of a variety of common molecular themes leading to the development of CHD. These include disruption of transcription factors and modification of chromatin, ligands, and signaling receptors [45].

## 8.7 The noncoding regulatory genome in congenital heart disease: microRNAs and circular RNAs

miRNAs are noncoding RNA strands that have emerged as important regulators in cardiac morphogenesis and myocardial cell growth. These small noncoding RNA strands (~16–24 nucleotides in length) regulate gene expression by binding to the 3'-untranslated region of target mRNA and subsequently affect biological functions at posttranscriptional and translational levels [46]. Zhu identified upregulation of four miRNAs in pregnant women with fetal CHD [47]. Similarly, Li found eight miRNAs to be dysregulated in patients with a ventricular septal defect in comparison to controls. Using gene-ontology analysis, the authors predicted that the principle target genes were associated with development of the right ventricle. Thus NOTCH1, HAND1, ZFPM2, and GATA3 were predicted as targets of hsa-let-7e-5p, miRNA-222-3p, and miRNA-433 [48]. Other groups have demonstrated that the NOTCH pathway is particularly susceptible to influence from miRNAs. Thus Wang demonstrated that, in comparison with a control group, miRNA-375 overexpression inhibited proliferation and promoted apoptosis in pluripotent cells (P19 cells), and associated miRNAs and proteins (in the NOTCH-signaling pathway) were decreased during differentiation. This confirmed that miRNA-375 can disrupt cardiomyocyte differentiation via modulation of the NOTCH-signaling pathway [49]. Liu demonstrated that upregulation of miRNA-1 has the opposite effect and promotes myocardial cell proliferation while suppressing apoptosis during heart development [50]. In a similar manner, upregulation of the NOTCH pathway is associated with ventricular septal defect development, while downregulation of the NOTCH pathway is associated with a reduced risk of developing HLHS [51,52].

Other workers have shown that miRNAs can influence the clinical consequences of an underlying congenital heart defect. Thus Huang et al. found that miRNA-184 is inhibited in children with cyanotic CHD, which they concluded was an adaptive mechanism to decrease the physiological myocardial cell proliferation that usually occurs as a response to chronic hypoxia [53]. Zhang recently demonstrated that miRNA-182 acts in the same way [54]. This opens up potential new therapeutic strategies to reduce the maladaptive responses to chronic hypoxia, which can occur in the cyanotic patient.

**TABLE 8.3** Examples of microRNAs (miRNAs) involved in congenital heart disease development (human studies).

miRNAs involved	Change	Congenital heart defect	Reference
miR-196a2	Upregulation	Cardiac septal defects via HOXB8 Sonic hedgehog	Xu [62]
miR-19b, miR-22, miR-29c, and miR-375	Upregulation	Atrial septal defect, ventricular septal defect, and tetralogy of Fallot	Zhu [47]
36 Differentially expressed miRNAs assessed	15 Upregulated 21 Downregulated	Ventricular septal defect	Li [48]
miR-375	Upregulation (NOTCH pathway)	Ventricular septal defect	Wang [49]
hsa-let-7a, hsa-let-7b, and hsa-miR-486	Upregulation	Atrial septal defect	Song [52]
miR-196a2	Varies with SNP	Atrial septal defect	Yu [58]
miR-184	Downregulation	Cyanotic congenital heart disease	Huang [53]
miR-182	Downregulation	Cyanotic congenital heart disease	Zhang [54]
miR-19a	Upregulation	Marker of pulmonary hypertension in congenital heart defects	Chen [63]
miR-592	Downregulation (NOTCH pathway)	Protects against development of hypoplastic left-heart syndrome	Pang [64]

SNP, Single-nucleotide polymorphism.

Many miRNAs identified in the fetus with CHD are of placental origin. Placental miRNAs can be detected in maternal blood, and the assessment of miRNAs in maternal blood has been proposed as a noninvasive method of identifying fetal CHD [55]. However, many effective methods for detecting fetal cardiac defects are already available (e.g., fetal echocardiography and ultrasound evaluation of nuchal translucency), and perhaps of more importance is the improved understanding that an evaluation of miRNAs provides in our attempts to understand the molecular mechanisms, which result in the development of congenital heart defects.

It remains unclear whether miRNA dysregulation is a cause as well as a consequence of the development of CHD. While more research is needed, the manipulation of miRNAs and their target gene expression using tools, such as chemically modified anti-miR oligonucleotides or miR-sponges, offer the potential to minimize fibrosis and hypertrophy or even enhance angiogenesis and improve clinical outcome in the future [56]. Hoelscher reviewed the interaction between miRNAs and CHD with particular emphasis on Holt–Oram syndrome, ventricular septal defects, TOF, and HLHS [56]. This paper concentrates on data obtained using animal models of CHD but there is increasing evidence that miRNAs may have a major influence in the development of CHD from studies carried out directly in man [57] (Table 8.3).

### 8.7.1 Congenital heart disease and single-nucleotide polymorphisms

A relatively new area of research is the assessment of the role of SNPs on miRNA expression. It is known that SNPs can alter miRNA expression. Yu et al. demonstrated that the rs11614913 (T > C) SNP of miR-196a2 is associated with the presence of an ASD but the homozygous CC variant is associated with a significantly reduced risk of an ASD [58]. A similar research development is the study of circular RNAs (circRNAs) and their influence on cardiac development. circRNAs are usually formed by alternative splicing of pre-mRNA such that the 3' and 5' ends are covalently linked. It is now appreciated that circRNAs are both widespread throughout and abundant within tissues. Although the precise function of circRNA is unclear, Liu has demonstrated that there is a significant difference in circRNA expression in fetuses with a VSD in comparison to those with a normal heart [51], suggesting that circRNAs may play an important role in the development of phenotype in CHD.

If vascular endothelial growth factor (VEGF) production is disrupted during early cardiac development, abnormalities develop in the endocardial cushions with consequent development of CHD—usually atrioventricular septal defects. Vannay demonstrated that the presence of the VEGF allele +405C was more common in neonates



with CHD than in a control group (odds ratio 1.72 CI 1.32–2.26). Presence of the –460CT/ +405GG allele association, however, was more common in controls suggesting that these polymorphisms may play an important role in the development of CHD [59]. In a similar way, certain polymorphisms in the TBX1 gene are associated with an increased risk of developing conotruncal defects [60]. Endothelial nitric oxide synthase (eNOS) produces nitric oxide that plays an important role in cell growth, apoptosis, and vessel dilatation. Mice deficient in eNOS develop CHD. In humans the 894G > T polymorphism in the eNOS gene is associated with birth defects. In an intriguing study by Van Beynum the eNOS 894G > T polymorphism, on its own, was not associated with CHD except in combination with a strong history of maternal smoking [61]. This provides evidence of a possible gene–environment association between this polymorphism and maternal smoking in pregnancy.

## 8.8 Noncardiac congenital anomalies in congenital heart disease

Noncardiac congenital anomalies are common in association with CHD. Stoll reported 4005 cases with CHD of whom 1055 (26.3%) had associated major noncardiac anomalies [27]. This included 99 (2.5%) with nonchromosomal-dysmorphic conditions. The majority of the nonchromosomal-dysmorphic conditions did not represent a recognizable syndrome with the exception of VACTERL syndrome (23%), Noonan syndrome (8%), fetal alcohol syndrome (7%), and skeletal dysplasias (6%). A further 602 cases had nonsyndromic, nonchromosomal multiple congenital anomalies of which anomalies in the urinary tract, the musculoskeletal, the digestive, and the central nervous systems were the most common [27].

The presence of noncongenital anomalies in CHD can illuminate some of the molecular mechanisms responsible for the underlying CHD. An example of this is the pleiotropic condition known as VACTERL syndrome that includes vertebral, anorectal, cardiac, tracheoesophageal, renal, and limb malformations (OMIM 192350). VACTERL syndrome was first described in 1973 by Quan and Smith [65]. Cardiac defects are frequently identified of which ventricular septal defects predominate occurring in 30% of cases [66]. In 2001 Kim et al. carried out a mutant analysis of GLI genes that encode transcription factors mediating Sonic hedgehog (Shh) signal transduction. They found that defective Shh signaling in the mouse led to developmental anomalies similar to those found in VACTERL. In 2008 Szumska identified nonsynonymous mutations in the proprotein convertase gene PCSK5 in patients with VACTERL [67]. This mutation (C470R) ablates a disulfide bond and blocks export from the endoplasmic reticulum and proprotein convertase activity. Although other mechanisms have now been identified as a cause for VACTERL [68], these genotype–phenotype correlations demonstrate that an evaluation of the molecular etiology of noncongenital cardiac anomalies in CHD may provide insights into the genetic causes of the complete spectrum of cardiovascular developmental disorder.

## 8.9 The epigenome and congenital heart disease

Epigenetics refers to the mechanisms that modify gene expression without changing the underlying nucleotide sequence. This can be thought of as a change in phenotype without any corresponding change in genotype. As progenitor cells differentiate, epigenetic mechanisms maintain stable gene expression while being both dynamic and subject to external stimuli. It is epigenetic mechanism that helps in explaining genetic heterogeneity and this is of particular importance in cardiovascular developmental disorders. One of the most common epigenetic processes occurs through RNA-based regulation. This has already been discussed in the section on miRNAs. However, epigenetic modifications can also occur through DNA methylation and histone modification.

### 8.9.1 DNA methylation

In DNA methylation a methyl group is covalently added to the five position of a cytosine nucleotide preceding a guanosine nucleotide to create 5-methylcytosine. This is known as a CpG dinucleotide. Most gene-promoter regions are located within the so-called CpG islands. CpG islands are DNA segments of >200 base pairs enriched with CpG dinucleotides (>60%). Methylation of CpG islands in the promoter region of genes results in decreased gene expression. DNA methylation can also occur within the gene body itself in regions known as CpG shores. DNA methylation in these regions can lead to varying degrees of gene regulation and transcription, not necessarily abolished gene expression [69,70].



Zaidi compared the incidence of de novo mutations in 362 severe CHD cases, and 264 controls by analyzing exome sequencing of parent–offspring trios [71]. The authors reported a marked excess of de novo mutations in genes involved in the production, removal, or reading of histone 3 lysine 4 (H3K4) methylation, or ubiquitination of H2BK120, which is required for H3K4 methylation. They also identified two de novo mutations in SMAD2, which is a regulator of H3K27 methylation in the embryonic left–right organizer. The combination of both activating (H3K4 methylation) and inactivating (H3K27 methylation) chromatin points to the importance of DNA methylation as a factor in the molecular etiology of CHD [72].

### 8.9.2 Histone modification

Guo carried out a whole-exome study of 184 individuals with 22q11 deletion syndrome of whom 89 had severe CHD [73]. In seven subjects, three genes (*JMJD1C*, *RREB1*, and *SEC24C*) were found, which had predicted deleterious single-nucleotide variations. *JMJD1C* and *RREB1* are involved in chromatin modification, so the authors investigated other histone modification genes. A further 18 subjects (20%) had predicted deleterious single-nucleotide variations in four genes (*JMJD1C*, *RREB1*, *MINA*, and *KDM7A*), which are involved in histone demethylation. These changes were not found in control subjects suggesting that histone modification may increase the risk of developing CHD in the presence of the 22q11 deletion [73].

## 8.10 Future considerations

Advances in genomics combined with the ready availability of NGS have improved our understanding of the changes in molecular physiology, which cause CHD. Multiple molecular pathways have been identified that regulate myocardial cell, tissue, and organ development. These complex and dynamic pathways are essential for the development of mature cardiovascular structure and function. An improved understanding of gene–environment interaction in these molecular events has the potential to facilitate new strategies for primary and secondary prevention of CHD. It is important to educate the cardiology community in these advances. The American Heart Association has responded by producing a clear scientific statement that describes the genetic causes of CHD and some of the molecular mechanisms involved [74]. The stated aim of this document is to facilitate comprehensive, interdisciplinary care for those with CHD. This is a laudable aim that offers much hope for future generations who are at risk of congenital cardiac malformations.

## References

- [1] Siffel C, Riehle-Colarusso T, Oster ME, Correa A. Survival of children with hypoplastic left heart syndrome. *Pediatrics* 2015;136(4):e864–70.
- [2] Stensen N. Embryo monstro affinis Parisiis dissectus. *Acta Med Philos Hafneincia* 1671;i:202–3.
- [3] van der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol* 2011;58(21):2241–7.
- [4] Dolk H, Loane M, Garne E, European Surveillance of Congenital Anomalies Working Group. Congenital heart defects in Europe: prevalence and perinatal mortality, 2000 to 2005. *Circulation* 2011;123(8):841–9.
- [5] Gross RE, Hubbard JP. Surgical ligation of a patent ductus arteriosus: report of first successful case. *JAMA* 1939;112:729.
- [6] Lillehei CW, Cohen M, Warden HE, Read RC, Aust JB, Dewall RA, et al. Direct vision intracardiac surgical correction of the tetralogy of Fallot, pentalogy of Fallot, and pulmonary atresia defects; report of first ten cases. *Ann Surg* 1955;142(3):418–42.
- [7] Kempny A, Dimopoulos K, Uebing A, Diller GP, Rosendahl U, Belitsis G, et al. Outcome of cardiac surgery in patients with congenital heart disease in England between 1997 and 2015. *PLoS One* 2017;12(6):e0178963.
- [8] Gittenberger-de Groot AC, Bartelings MM, Deruiter MC, Poelmann RE. Basics of cardiac development for the understanding of congenital heart malformations. *Pediatr Res* 2005;57(2):169–76.
- [9] Mitchell ME, Sander TL, Klinkner DB, Tomita-Mitchell A. The molecular basis of congenital heart disease. *Semin Thorac Cardiovasc Surg* 2007;19(3):228–37.
- [10] Blue GM, Kirk EP, Giannoulitou E, Sholler GF, Dunwoodie SL, Harvey RP, et al. Advances in the genetics of congenital heart disease: a clinician's guide. *J Am Coll Cardiol* 2017;69(7):859–70.
- [11] Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, et al. Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* 1998;281(5373):108–11.
- [12] Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soultis J, et al. Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt–Oram syndrome. *Nat Genet* 1997;15(1):30–5.
- [13] Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, et al. Mutations in NOTCH1 cause aortic valve disease. *Nature* 2005;437(7056):270–4.

- [14] Sultana N, Nag K, Hoshijima K, Laird DW, Kawakami A, Hirose S. Zebrafish early cardiac connexin, Cx36.7/Ecx, regulates myofibril orientation and heart morphogenesis by establishing Nkx2.5 expression. *Proc Natl Acad Sci USA* 2008;105(12):4763–8.
- [15] Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med* 2010;363(2):166–76.
- [16] Cordell HJ, Topf A, Mamasoula C, Postma AV, Benthall J, Zelenika D, et al. Genome-wide association study identifies loci on 12q24 and 13q32 associated with tetralogy of Fallot. *Hum Mol Genet* 2013;22(7):1473–81.
- [17] Agopian AJ, Mitchell LE, Glessner J, Bhalla AD, Sewda A, Hakonarson H, et al. Genome-wide association study of maternal and inherited loci for conotruncal heart defects. *PLoS One* 2014;9(5):e96057.
- [18] Cordell HJ, Benthall J, Topf A, Zelenika D, Heath S, Mamasoula C, et al. Genome-wide association study of multiple congenital heart disease phenotypes identifies a susceptibility locus for atrial septal defect at chromosome 4p16. *Nat Genet* 2013;45(7):822–4.
- [19] Mamasoula C, Prentice RR, Pierscionek T, Pangilinan F, Mills JL, Druschel C, et al. Association between C677T polymorphism of methylene tetrahydrofolate reductase and congenital heart disease: meta-analysis of 7697 cases and 13,125 controls. *Circ Cardiovasc Genet* 2013;6(4):347–53.
- [20] Adams DR, Eng CM. Next-generation sequencing to diagnose suspected genetic disorders. *N Engl J Med* 2018;379(14):1353–62.
- [21] Pawlak M, Niescierowicz K, Winata CL. Decoding the heart through next generation sequencing approaches. *Genes (Basel)* 2018;9(6).
- [22] Bonachea EM, Zender G, White P, Corsmeier D, Newsom D, Fitzgerald-Butt S, et al. Use of a targeted, combinatorial next-generation sequencing approach for the study of bicuspid aortic valve. *BMC Med Genomics* 2014;7:56.
- [23] Liu X, Yagi H, Saeed S, Bais AS, Gabriel GC, Chen Z, et al. The complex genetics of hypoplastic left heart syndrome. *Nat Genet* 2017;49(7):1152–9.
- [24] Pfitzer C, Helm PC, Rosenthal LM, Berger F, Bauer UMM, Schmitt KR. Dynamics in prevalence of Down syndrome in children with congenital heart disease. *Eur J Pediatr* 2018;177(1):107–15.
- [25] Freeman SB, Bean LH, Allen EG, Tinker SW, Locke AE, Druschel C, et al. Ethnicity, sex, and the incidence of congenital heart defects: a report from the National Down Syndrome Project. *Genet Med* 2008;10(3):173–80.
- [26] Pelleri MC, Gennari E, Locatelli C, Piovesan A, Caracausi M, Antonaros F, et al. Genotype-phenotype correlation for congenital heart disease in Down syndrome through analysis of partial trisomy 21 cases. *Genomics* 2017;109(5–6):391–400.
- [27] Stoll C, Dott B, Alembik Y, Roth MP. Associated noncardiac congenital anomalies among cases with congenital heart defects. *Eur J Med Genet* 2015;58(2):75–85.
- [28] Botto LD, May K, Fernhoff PM, Correa A, Coleman K, Rasmussen SA, et al. A population-based study of the 22q11.2 deletion: phenotype, incidence, and contribution to major birth defects in the population. *Pediatrics* 2003;112(1 Pt 1):101–7.
- [29] Digilio MC, Angioni A, De Santis M, Lombardo A, Giannotti A, Dallapiccola B, et al. Spectrum of clinical variability in familial deletion 22q11.2: from full manifestation to extremely mild clinical anomalies. *Clin Genet* 2003;63(4):308–13.
- [30] Morrow BE, McDonald-McGinn DM, Emanuel BS, Vermeesch JR, Scambler PJ. Molecular genetics of 22q11.2 deletion syndrome. *Am J Med Genet, A* 2018;176(10):2070–81.
- [31] Xing HJ, Li YJ, Ma QM, Wang AM, Wang JL, Sun M, et al. Identification of microRNAs present in congenital heart disease associated copy number variants. *Eur Rev Med Pharmacol Sci* 2013;17(15):2114–20.
- [32] Tartaglia M, Mehler EL, Goldberg R, Zampino G, Brunner HG, Kremer H, et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat Genet* 2001;29(4):465–8.
- [33] Fragale A, Tartaglia M, Wu J, Gelb BD. Noonan syndrome-associated SHP2/PTPN11 mutants cause EGF-dependent prolonged GAB1 binding and sustained ERK2/MAPK1 activation. *Hum Mutat* 2004;23(3):267–77.
- [34] Clark KL, Yutzey KE, Benson DW. Transcription factors and congenital heart defects. *Annu Rev Physiol* 2006;68:97–121.
- [35] Fan C, Chen Q, Wang QK. Functional role of transcriptional factor TBX5 in pre-mRNA splicing and Holt-Oram syndrome via association with SC35. *J Biol Chem* 2009;284(38):25653–63.
- [36] Takeuchi JK, Lou X, Alexander JM, Sugizaki H, Delgado-Olguin P, Holloway AK, et al. Chromatin remodelling complex dosage modulates transcription factor function in heart development. *Nat Commun* 2011;2:187.
- [37] Waldron L, Steimle JD, Greco TM, Gomez NC, Dorris KM, Kweon J, et al. The cardiac TBX5 interactome reveals a chromatin remodeling network essential for cardiac septation. *Dev Cell* 2016;36(3):262–75.
- [38] Pagon RA, Graham Jr. JM, Zonana J, Yong SL. Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. *J Pediatr* 1981;99(2):223–7.
- [39] Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, et al. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* 2004;36(9):955–7.
- [40] Zentner GE, Hurd EA, Schnetz MP, Handoko L, Wang C, Wang Z, et al. CHD7 functions in the nucleolus as a positive regulator of ribosomal RNA biogenesis. *Hum Mol Genet* 2010;19(18):3491–501.
- [41] Ikeda Y, Hiroi Y, Hosoda T, Utsunomiya T, Matsuo S, Ito T, et al. Novel point mutation in the cardiac transcription factor CSX/NKX2.5 associated with congenital heart disease. *Circ J* 2002;66(6):561–3.
- [42] Ellesøe SG, Johansen MM, Bjerre JV, Hjortdal VE, Brunak S, Larsen LA. Familial atrial septal defect and sudden cardiac death: identification of a novel NKX2-5 mutation and a review of the literature. *Congenit Heart Dis* 2016;11(3):283–90.
- [43] Chen Y, Han ZQ, Yan WD, Tang CZ, Xie JY, Chen H, et al. A novel mutation in GATA4 gene associated with dominant inherited familial atrial septal defect. *J Thorac Cardiovasc Surg* 2010;140(3):684–7.
- [44] Su W, Zhu P, Wang R, Wu Q, Wang M, Zhang X, et al. Congenital heart diseases and their association with the variant distribution features on susceptibility genes. *Clin Genet* 2017;91(3):349–54.
- [45] Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: the glass half empty. *Circ Res* 2013;112(4):707–20.
- [46] Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 2010;466(7308):835–40.
- [47] Zhu S, Cao L, Zhu J, Kong L, Jin J, Qian L, et al. Identification of maternal serum microRNAs as novel non-invasive biomarkers for pre-natal detection of fetal congenital heart defects. *Clin Chim Acta* 2013;424:66–72.

- [48] Li D, Ji L, Liu L, Liu Y, Hou H, Yu K, et al. Characterization of circulating microRNA expression in patients with a ventricular septal defect. *PLoS One* 2014;9(8):e106318.
- [49] Wang L, Song G, Liu M, Chen B, Chen Y, Shen Y, et al. MicroRNA-375 overexpression influences P19 cell proliferation, apoptosis and differentiation through the Notch signaling pathway. *Int J Mol Med* 2016;37(1):47–55.
- [50] Liu L, Yuan Y, He X, Xia X, Mo X. MicroRNA-1 upregulation promotes myocardiocyte proliferation and suppresses apoptosis during heart development. *Mol Med Rep* 2017;15(5):2837–42.
- [51] Liu H, Hu Y, Zhuang B, Yin J, Chen XH, Wang J, et al. Differential expression of CircRNAs in embryonic heart tissue associated with ventricular septal defect. *Int J Med Sci* 2018;15(7):703–12.
- [52] Song Y, Higgins H, Guo J, Harrison K, Schultz EN, Hales BJ, et al. Clinical significance of circulating microRNAs as markers in detecting and predicting congenital heart defects in children. *J Transl Med* 2018;16(1):42.
- [53] Huang J, Li X, Li H, Su Z, Wang J, Zhang H. Down-regulation of microRNA-184 contributes to the development of cyanotic congenital heart diseases. *Int J Clin Exp Pathol* 2015;8(11):14221–7.
- [54] Zhang Y, Peng B, Han Y. MiR-182 alleviates the development of cyanotic congenital heart disease by suppressing HES1. *Eur J Pharmacol* 2018;836:18–24.
- [55] Yu Z, Han S, Hu P, Zhu C, Wang X, Qian L, et al. Potential role of maternal serum microRNAs as a biomarker for fetal congenital heart defects. *Med Hypotheses* 2011;76(3):424–6.
- [56] Hoelscher SC, Doppler SA, Dressen M, Lahm H, Lange R, Krane M. MicroRNAs: pleiotropic players in congenital heart disease and regeneration. *J Thorac Dis* 2017;9(Suppl. 1):S64–81.
- [57] Smith T, Rajakaruna C, Caputo M, Emanuelli C. MicroRNAs in congenital heart disease. *Ann Transl Med* 2015;3(21):333.
- [58] Yu K, Ji Y, Wang H, Xuan QK, Li BB, Xiao JJ, et al. Association of miR-196a2, miR-27a, and miR-499 polymorphisms with isolated congenital heart disease in a Chinese population. *Genet Mol Res* 2016;15(4).
- [59] Vannay A, Vasarhelyi B, Kornyei M, Treszl A, Kozma G, Gyorffy B, et al. Single-nucleotide polymorphisms of VEGF gene are associated with risk of congenital valvuloseptal heart defects. *Am Heart J* 2006;151(4):878–81.
- [60] Han XM, Lou Y, Zhu XY, Hu XF, Pang WY, Sun ZJ, et al. [Single nucleotide polymorphism and haplotype in TBX1 gene of patients with conotruncal defects: analysis of 130 cases]. *Zhonghua yi xue za zhi* 2006;86(22):1553–7.
- [61] van Beynum IM, Mooij C, Kapusta L, Heil S, den Heijer M, Blom HJ. Common 894G > T single nucleotide polymorphism in the gene coding for endothelial nitric oxide synthase (eNOS) and risk of congenital heart defects. *Clin Chem Lab Med* 2008;46(10):1369–75.
- [62] Xu J, Hu Z, Xu Z, Gu H, Yi L, Cao H, et al. Functional variant in microRNA-196a2 contributes to the susceptibility of congenital heart disease in a Chinese population. *Hum Mutat* 2009;30(8):1231–6.
- [63] Chen W, Li S. Circulating microRNA as a novel biomarker for pulmonary arterial hypertension due to congenital heart disease. *Pediatr Cardiol* 2017;38(1):86–94.
- [64] Pang XF, Lin X, Du JJ, Zeng DY. Downregulation of microRNA-592 protects mice from hypoplastic heart and congenital heart disease by inhibition of the Notch signaling pathway through upregulating KCTD10. *J Cell Physiol* 2019;234.
- [65] Quan L, Smith DW. The VATER association. Vertebral defects, Anal atresia, T-E fistula with esophageal atresia, Radial and Renal dysplasia: a spectrum of associated defects. *J Pediatr* 1973;82(1):104–7.
- [66] Khoury MJ, Cordero JF, Greenberg F, James LM, Erickson JD. A population study of the VACTERL association: evidence for its etiologic heterogeneity. *Pediatrics* 1983;71(5):815–20.
- [67] Szumska D, Pieleś G, Essalmani R, Bilski M, Mesnard D, Kaur K, et al. VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase Pcsk5. *Genes Dev* 2008;22(11):1465–77.
- [68] Reutter H, Hilger AC, Hildebrandt F, Ludwig M. Underlying genetic factors of the VATER/VACTERL association with special emphasis on the “Renal” phenotype. *Pediatr Nephrol* 2016;31(11):2025–33.
- [69] Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;13(7):484–92.
- [70] Osborne-Majnik A, Fu Q, Lane RH. Epigenetic mechanisms in fetal origins of health and disease. *Clin Obstet Gynecol* 2013;56(3):622–32.
- [71] Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 2013;498(7453):220–3.
- [72] Gupta RM. Exome sequencing in congenital heart disease points to importance of DNA methylation. *Circ Cardiovasc Genet* 2013;6(5):522.
- [73] Guo T, Chung JH, Wang T, McDonald-McGinn DM, Kates WR, Hawula W, et al. Histone modifier genes alter conotruncal heart phenotypes in 22q11.2 deletion syndrome. *Am J Hum Genet* 2015;97(6):869–77.
- [74] Pierpont ME, Brueckner M, Chung WK, Garg V, Lacro RV, McGuire AL, et al. Genetic basis for congenital heart disease: revisited: a scientific statement from the American heart association. *Circulation* 2018;138(21):e653–711.



# Channelopathies in clinical medicine—cardiac arrhythmias

Julian O.M. Ormerod<sup>1,2</sup> and Elijah R. Behr<sup>3</sup>

<sup>1</sup>Department of Cardiology, Milton Keynes General Hospital, Milton Keynes, United Kingdom <sup>2</sup>Department of Cardiovascular Medicine, University of Oxford, United Kingdom <sup>3</sup>Department of Cardiovascular Medicine, St. George's Hospital Medical School, London, United Kingdom

## 9.1 Introduction

Inherited arrhythmia syndromes are rare, but collectively they are a major cause of loss of life in otherwise healthy young people. Several syndromes are described, and the pathogenic mechanisms and molecular genetics are increasingly well understood. Most of them affect ion currents and are termed “channelopathies.”

## 9.2 Clinical cases

### 9.2.1 Case 1

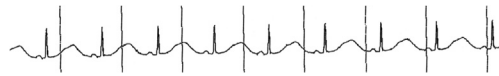
A 32-year-old woman presents to the emergency department with an increased frequency of “seizures.” She was diagnosed at the age of 15 with idiopathic epilepsy, and her attacks are characterized by unheralded loss of consciousness and postural tone, with pallor and (on occasions) generalized convulsions. She had tried multiple combinations of antiepileptic medications with no success. Three weeks earlier she had given birth to her first child. Blood testing revealed hypokalemia ( $K^+$ —2.9 mmol/L), and the rhythm strip (lead II) of her electrocardiogram (ECG) is shown in Fig. 9.1.

She subsequently suffered syncope with convulsions in the department, and the ECG showed ventricular fibrillation which was converted to sinus rhythm with a single 200 J biphasic shock. On further questioning, it was conveyed that her maternal aunt died suddenly in the postpartum period, and her mother suffered from multiple episodes of syncope before dying suddenly in her early 40s. She was treated with intravenous  $K^+$  supplementation and monitored on the acute cardiac unit. Echocardiography and CT coronary angiography were normal, and a clinical diagnosis of long QT was made. She went on to receive a dual chamber transvenous defibrillator and was started on nadolol 40 mg BD (a nonselective  $\beta$ -blocker). Subsequent genetic analysis revealed a missense mutation in a pore-forming region of *KCNH2*, confirming a diagnosis of long QT syndrome (LQTS) type 2 and cascade testing of family members was started.

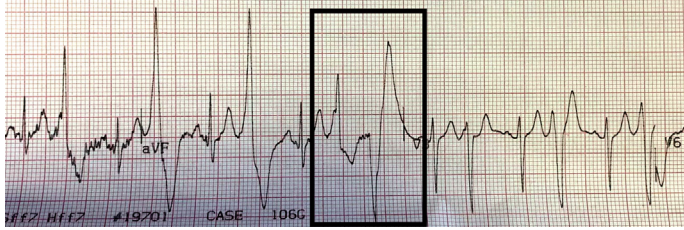
### 9.2.2 Case 2

A 13-year-old boy presents to his family doctor with dizzy spells and blackouts that tend to occur when he plays football. He is referred to a pediatric cardiologist who performs a 12-lead ECG and echocardiogram, which are both normal. An exercise tolerance test is performed, which shows ventricular ectopy at peak exercise, including a single bidirectional couplet (shown in Fig. 9.2).

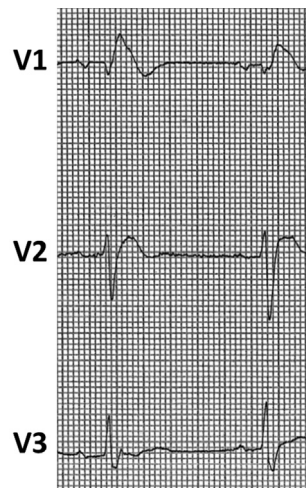




**FIGURE 9.1** ECG strip showing prolonged QT interval.



**FIGURE 9.2** ECG showing frequent ectopy with a single bidirectional couplet (highlighted with black box) at peak exercise.



**FIGURE 9.3** A 12-lead ECG showing type 1 Brugada pattern. ECG, Electrocardiogram.

Genetic testing revealed a mutation in *RYR2*, which had been previously described in multiple families with catecholaminergic polymorphic ventricular tachycardia (CPVT), confirming this diagnosis. Both parents were tested, showing this was a *de novo* mutation in this family, and therefore with a low risk of affecting his siblings.  $\beta$ -blockers were started and uptitrated to the maximum tolerated dose. He had no further syncope for 3 years and then suffered three further episodes in rapid succession during an argument with his parents. He was referred for left cardiac sympathetic denervation (LCSD).

### 9.2.3 Case 3

A 44-year-old man presented to the emergency department by ambulance. His wife woke up from sleep to see him breathing very abnormally and could not feel a pulse. He awoke spontaneously and was mildly confused, reporting that he had suffered a nightmare. Several years before, his older brother had died in his sleep and so the patient's wife convinced him to call an ambulance. A 12-lead ECG, taken on arrival, showed coved ST elevation (greater than 2 mm) terminating in a negative T wave in lead V1, which was recognized as a type 1 Brugada ECG pattern (Fig. 9.3).

The patient was referred to cardiology, and on the basis of a clear history of nocturnal agonal breathing (a surrogate for cardiac arrest), he was given a diagnosis of Brugada syndrome (BrS) and fitted with a subcutaneous defibrillator. He was advised to avoid certain drugs and to treat fever promptly. Genetic testing was performed,



but it was negative, and clinical screening was cascaded to first-degree relatives. He was asymptomatic for several months before suffering an appropriate shock for ventricular fibrillation in the early hours of the morning. He was started on quinidine with good effect.

## 9.3 Molecular systems underpinning the clinical scenario

### 9.3.1 Overview of the cardiac action potential

The resting membrane potential of cardiomyocytes is approximately  $-85$  to  $-95$  mV. It is largely determined by the  $K^+$  membrane potential (as the membrane is relatively less permeable to other ions) and is set up by several electrochemical exchange pumps and channels leading to net differences in ion concentrations between the intracellular and extracellular spaces. The resting potential and the exact shape of the cardiac action potential vary throughout the different tissues of the heart (e.g., sinoatrial pacemaker cells lack the electrical plateau), but general principles of cardiac conduction are explained here. Under normal circumstances the initial impulse occurs in a pacemaker cell, typically in the sinoatrial node (due to a slow depolarizing “funny” current,  $i_f$ ). Electrical impulses are passed from cell to cell through gap junctions, and this depolarizing impulse triggers an action potential in subsequent cells.

The cardiac action potential begins with rapid depolarization (phase 0), mediated by opening of  $Na^+$  channels and flow of  $Na^+$  into the cell. Ion currents are conventionally named with an initial letter “i”, and hence, this is termed  $i_{Na}$ . Each current may be mediated by multiple voltage-gated channel types, the details of which are beyond the scope of this chapter but are reviewed here (molecular physiology of cardiac repolarization). Following this very rapid depolarization, early  $i_{Na}$ -carrying channels are rapidly inactivated, and there is a partial repolarization caused by the transient outward current,  $i_{to}$ , mediated by  $K^+$  and (probably)  $Cl^-$  ions. This produces the characteristic notch of phase 1. The membrane potential reaches a plateau in phase 2, produced by a balance of inward  $Ca^{2+}$  flow ( $i_{Ca}$ ) through “L-type”  $Ca^{2+}$  channels in the membrane, late  $i_{Na}$  current, and outward “slow rectifying”  $K^+$  flow ( $i_{KS}$ ), which continues into phase 3. Repolarization takes place in phase 3, with inactivation of  $i_{Ca}$ , and activation of additional further repolarizing  $K^+$  currents, “rapid delayed rectifying”  $i_{Kr}$  and “inward rectifying” (previously known as “anomalous rectifying”)  $i_{K1}$ .

### 9.3.2 Relation of the action potential to the surface electrocardiogram

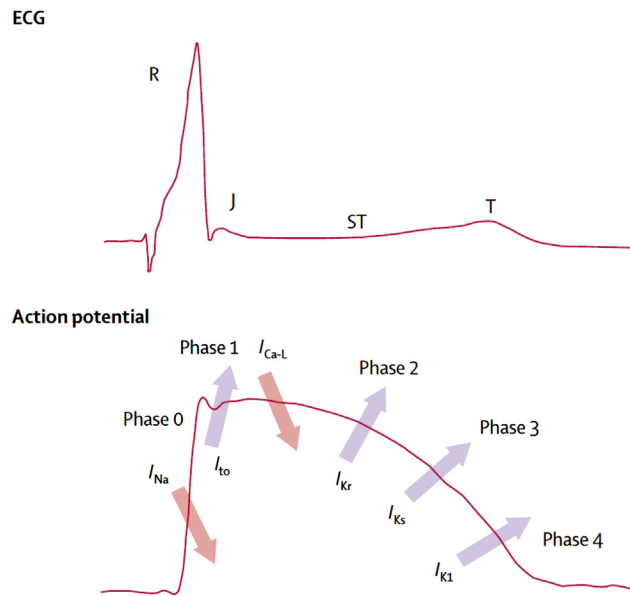
The cardiac action potential has been investigated elucidated using traditional patch clamp and sharp electrode techniques in single cells and myocardial wedge preparations. However, the main clinical tool used to investigate patients is the 12-lead surface ECG. The power of the ECG to diagnose arrhythmia conditions is augmented by various tools including provocation with drugs (adrenaline or sodium-channel blockers such as ajmaline) or exercise, signal averaging, and prolonged recording (e.g., Holter monitoring) (Fig. 9.4).

Each vector (or lead) of the surface ECG reflects the summation of many individual action potentials. It is predominantly influenced by atrial and ventricular myocardium, as the specialized conduction tissue is comparatively very small in mass. The p wave represents atrial depolarization, the QRS complex begins with depolarization of ventricular muscle, and its width reflects the time taken for all muscle to depolarize. The ST segment corresponds to the plateau phase and shifts from the baseline (i.e., ST elevation or depression) reflect voltage gradients from epicardium to endocardium (with the largest contribution from the middle or M cells) and are commonly caused by ischemia (which tends to affect endocardium first due to the nature of coronary blood supply) or infarction. The T wave represents the repolarization phase.

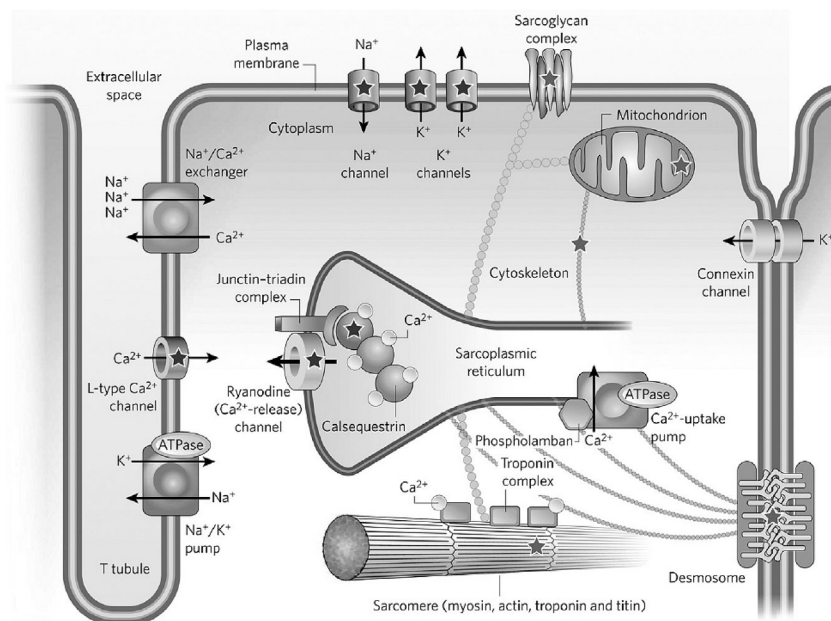
Importantly, the QT interval (measured between the beginning of the earliest Q wave and the end of the latest T wave) does not solely measure the repolarization phase of the action potential, but in fact it measures the entire of the depolarization–repolarization period. Slowed conduction in cardiomyopathy, for example, leads to QRS prolongation (e.g., left bundle branch block), which also therefore prolongs the QT interval. LQTS can only be diagnosed in the absence of structural heart disease or coronary artery disease.

### 9.3.3 The sarcoplasmic reticulum and excitation–contraction coupling

Fig. 9.5 shows sarcoplasmic reticulum (SR) and interaction with cytosolic sodium overload and effect on calcium.



**FIGURE 9.4** Correlation of ECG segments with repolarization and depolarization phases. Source: Reprinted from Morita H, Wu J, Zipes DP. The QT syndromes: long and short. *Lancet* 2008;372(9640):750–63. [https://doi.org/10.1016/S0140-6736\(08\)61307-0](https://doi.org/10.1016/S0140-6736(08)61307-0) [1] with permission.



**FIGURE 9.5** Source: Reprinted with permission from Watanabe H, Knollmann BC. Mechanism underlying catecholaminergic polymorphic ventricular tachycardia and approaches to therapy. *J Electrocardiol* 2011;44(6):650–5. <https://doi.org/10.1016/j.jelectrocard.2011.07.025> [2].

As mentioned previously, the cardiac action potential is conducted rapidly to every area of the ventricles, and then into muscle tissue. Cardiomyocytes have a lipid bilayer cell wall (the sarcolemma) and contain multiple myofibrils constructed from repeated basic contractile units known as sarcomeres and a specialized version of endoplasmic reticulum (SR). Sarcomeres contain several proteins, including myosin, actin, and tropomyosin. Mutations in sarcomeric proteins cause hypertrophic cardiomyopathy.

When an action potential arrives at the cardiomyocyte, voltage-gated L-type calcium channels in the sarcolemma allow influx of extracellular  $\text{Ca}^{2+}$  into the cytoplasmic space. This  $\text{Ca}^{2+}$  activates a second type of calcium channel in the membrane of the SR, known as the ryanodine receptor (RyR) for historical reasons, leading to

release of SR stores of  $\text{Ca}^{2+}$  (a process known as calcium-induced calcium release). The SR can store an extraordinary amount of  $\text{Ca}^{2+}$ ; both by physical separation of the cisternae from the cytosol by the SR membrane, and by the  $\text{Ca}^{2+}$ -binding capacity of a protein, calsequestrin. The released  $\text{Ca}^{2+}$  binds to tropomyosin, which changes shape and uncovers the binding sites on actin. The myosin heads act as an ATP-dependent molecular motor, ratcheting along the actin molecules and causing contraction.

### 9.3.4 Generation and propagation of clinical arrhythmias

Clinical arrhythmias are caused by three main mechanisms: automaticity (increased pacemaker activity), re-entry, and triggered activity (including early afterdepolarizations). The latter two are most important in inherited arrhythmia syndromes. These mechanisms underlie both atrial and ventricular arrhythmias.

#### 9.3.4.1 Increased automaticity

Ventricular myocytes do not usually exhibit pacemaker activity. With the phenomenon of increased automaticity, cells depolarize more frequently (and therefore may be a source of arrhythmias). However, in contrast to triggered activity (below) they still only depolarize once per action potential.

#### 9.3.4.2 Re-entry

Depolarization spreads rapidly through the specialized conduction tissue of the heart and into the muscle cells, causing coordinated contraction. Depolarization spreads from base to apex and from end to epicardium, and repolarization proceeds in the opposite sequence. However, due to rapidity of this process, relatively nearby cardiomyocytes depolarize and repolarize approximately at the same time and are then refractory to further impulses. If conduction into a given region of the ventricle is slow, and therefore depolarization is delayed, normal tissue may have a chance to recover and become excitable just as the impulse exits the abnormal region. This requires electrical isolation of the region (e.g., a scar) and unidirectional conduction so the wavefront does not extinguish itself. If these conditions are met, re-entry sets up a “circus rhythm,” which can be self-sustaining and (depending upon the length of each limb of the circuit) may be very rapid. The heart fills during diastole, and so extreme tachycardia prevents adequate filling. The cardiac output and blood pressure drops, leading to syncope and eventually death. Re-entry is likely to be the mechanism sustaining arrhythmia in several inherited arrhythmia conditions.

#### 9.3.4.3 Triggered activity

This is distinguishable from automaticity by the observation that in triggered activity cells may depolarize more than once per action potential. This phenomenon is caused by early or delayed afterdepolarizations. Early afterdepolarizations occur during phase 2 or 3 and are caused by the prolongation of repolarization due to decreased  $i_K$  or increased late  $i_{Na}$ ; this allows reactivation of L-type  $\text{Ca}^{2+}$  channels (and hence depolarization) and is likely to be an important mechanism of arrhythmogenesis in LQTS. Delayed afterdepolarizations occur during phase 4 and are caused by increased cytosolic  $\text{Ca}^{2+}$ . This was classically described in digoxin toxicity and is the mechanism of arrhythmogenesis in CPVT.

## 9.4 Overview of molecular biology and pathophysiology

The QT interval may be prolonged or shortened by perturbation of  $\text{K}^+$  ( $i_{Ks}$ ,  $i_{Kr}$ , and much less commonly  $i_{K1}$ ),  $\text{Na}^+$ , or (rarely)  $\text{Ca}^{2+}$  currents, leading to the LQTS and short QT syndrome (SQTS). BrS is thought to be caused by decreased early (phase 0)  $\text{Na}^+$  current or increased transient outward current,  $i_{to}$ . CPVT is caused by a different mechanism, the channels are known to be involved in the SR rather than the cell membrane, and increased intracellular  $\text{Ca}^{2+}$  is thought to be the cause of arrhythmia.

### 9.4.1 Mechanisms of channelopathy

The mechanisms of channelopathy are loss of function, decreased traffic to membrane, poison peptide, gain of function, and ankyrin (LQT4).

## 9.4.2 Long QT syndrome

### 9.4.2.1 Overview of long QT syndrome

Monogenic (as opposed to polygenic) LQTS is a rare inherited disease affecting approximately 1:2000 of the population. Autosomal-dominant forms are most common, but there are also very rare recessive forms. Mutations in many genes have been shown to cause monogenic LQTS and these are summarized in Table 9.1.

### 9.4.2.2 Long QT syndrome diagnosis

Congenital LQTS is generally diagnosed on the basis of the surface ECG, which shows a prolonged QT interval. This is the time between the start of the QRS complex and the end of the T wave. Measurement is complicated by several factors. It is not always clear exactly where the T wave ends, and convention is to measure where a line drawn against the steepest part of the downslope of the T wave crosses the baseline. This will seriously underestimate the true QT interval in some patients (e.g., Andersen–Tawil syndrome). U waves are positive deflections seen after the T wave in some individuals and are considered benign but distinguishing them from the very abnormal T wave morphology seen in some LQTS patients can be challenging. Finally, the normal QT interval shortens with increasing heart rate. Formulae are available to “correct” for this, the most commonly used being that of Bazett ( $QT_c = QT/\sqrt{RR}$ ), but these are very unreliable outside a relatively narrow range (e.g., 50–80 bpm).

The QT interval is variable across the population, and so simply being outside the normal range does not confer a diagnosis of LQTS. Diagnosis therefore depends upon the corrected QT interval (the  $QT_c$ ) alongside a number of other factors that increase the pretest probability of clinical disease. This can be expressed arithmetically by the Schwartz Score (Table 9.2).

Dominantly inherited LQTS without extracardiac features was previously known as Romano–Ward syndrome, but now the molecular basis is better understood, the specific genetic lesion is often known. Genetic testing discovers a pathogenic variant in 80%–90% of individuals tested. Some of the remainder—perhaps as much

**TABLE 9.1** Classification of the genetic causes of long QT syndrome.

	Gene	Protein	Current	Clinical features
LQT1/JLN1	KCNQ1	K <sub>v</sub> 7.1	↓ <i>i</i> <sub>Ks</sub>	JLN1—long QT and sensorineural deafness
LQT2	KCNH2	hERG	↓ <i>i</i> <sub>Kr</sub>	
LQT3	SCN5A	Na <sub>v</sub> 1.5	↑ <i>i</i> <sub>Na</sub>	
LQT4	ANK2	Ankyrin-b	Multiple	
LQT5/JLN2	KCNE1	MinK	↓ <i>i</i> <sub>Ks</sub>	JLN2—long QT and sensorineural deafness
LQT6	KCNE2	MiRP1	↓ <i>i</i> <sub>Kr</sub>	Drug-induced QT prolongation
LQT7	KCNJ2	K <sub>ir</sub> 2.1	↓ <i>i</i> <sub>K1</sub>	Andersen–Tawil syndrome (some cases)
LQT8	CACNA1C	Ca <sub>v</sub> 1.2	↑ <i>i</i> <sub>Ca</sub>	Timothy syndrome (some cases)
LQT9	CAV3	Caveolin-3	↑ <i>i</i> <sub>Na</sub>	Similar to LQT3
LQT10	SCN4B	Na <sub>v</sub> 1.5	↑ <i>i</i> <sub>Na</sub>	Similar to LQT3
LQT11	AKAP9	Kinase-A anchor protein-9	↓ <i>i</i> <sub>Ks</sub>	Similar to LQT1
LQT12	SNTA1	α1-syntrophin	↑ <i>i</i> <sub>Na</sub>	Similar to LQT3
LQT13	KCNJ5	K <sub>ir</sub> 3.5	↓ <i>i</i> <sub>KATP</sub>	
LQT14	CALM1	Calmodulin	Multiple	
LQT15	CALM2	Calmodulin	Multiple	
LQT16	CALM3	Calmodulin	Multiple	
	SCN1B	Na <sub>v</sub> 1.5	↑ <i>i</i> <sub>Na</sub>	
	RYR2	RyR2	↑ <i>i</i> <sub>Ca</sub>	
	TRDN	Triadin	↑ <i>i</i> <sub>Ca</sub>	

**TABLE 9.2** The Schwartz score can be used to estimate the likelihood of congenital long QT syndrome (LQTS) in a given patient.

		Score
<b>ECG findings</b>		
A.	QT <sub>c</sub> > 480 ms	3
	QT <sub>c</sub> 460–479 ms	2
	QT <sub>c</sub> 450–459 ms (in males)	1
B.	QT <sub>c</sub> at fourth minute of recovery from exercise stress test ≥ 480 ms	1
C.	Torsades-de-pointes <sup>a</sup>	2
D.	T wave alternans	1
E.	Notched T wave in three leads	1
F.	Resting heart rate below the second percentile for age	0.5
<b>Clinical history</b>		
A.	Syncope with stress <sup>a</sup>	2
	Syncope without stress	1
B.	Congenital deafness	0.5
<b>Family history</b>		
A.	Family members with definite LQTS <sup>b</sup>	1
B.	Unexplained sudden cardiac death below age 30 among immediate family members <sup>b</sup>	0.5

<sup>a</sup>Cannot score for both torsades-de-pointes and syncope.<sup>b</sup>Cannot be the same family member.

Notes: ≤ 1 point: low probability of LQTS, 1.5–3 points: intermediate probability of LQTS, ≥ 3.5 points: high probability; QT corrected by Bazett's formula; patient not on QT-prolonging drugs. ECG, Electrocardiogram.

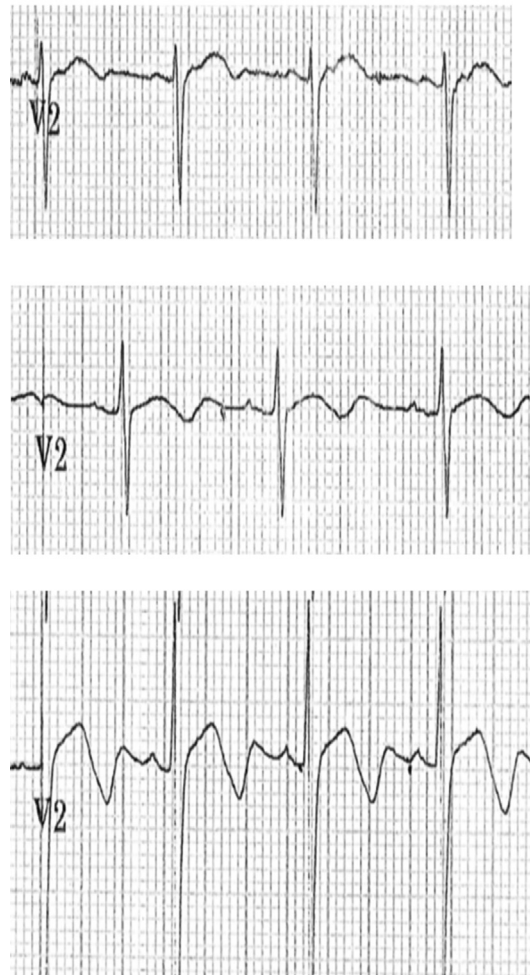
Adapted from Schwartz PJ, Crotti L. QTc behavior during exercise and genetic testing for the long-QT syndrome. *Circulation*. 2011;124:2181–2184 [3].**FIGURE 9.6** Distinctive ECG patterns in LQT1–3 and LQT7 (Anderson–Tawil syndrome). ECG, Electrocardiogram. Source: Reprinted from Morita H, Wu J, Zipes DP. The QT syndromes: long and short. *Lancet* 2008;372(9640):750–63. [https://doi.org/10.1016/S0140-6736\(08\)61307-0](https://doi.org/10.1016/S0140-6736(08)61307-0) with permission.

as 3%—may be caused by an abnormal copy number of LQTS-associated genes [4], and this will often be missed unless looked for specifically. LQTS types 1–3 make up around 95% of positively genotyped cases. Consequently, the clinical features of these conditions are best studied. In addition, various clinical syndromes with prolonged QT interval, arrhythmias, and associated extracardiac features are also described (Figs. 9.6 and 9.7).

#### 9.4.2.2.1 Long QT syndrome type 1

Loss-of-function mutations in *KCNQ1*, which encodes the  $\alpha$ -subunit of the  $K_v7.1$  channel, cause LQT1. This protein assembles as a tetramer and carries the outward slow delayed rectifying  $K^+$  current ( $i_{Ks}$ ), and deficiency prolongs the repolarization phase of the action potential. The ECG tends to show tall, broad-based T waves. The proteins encoded by *KCNE1* (LQT5) and *AKAP9* (LQT11) are associated with the  $K_v7.1$  channel; and hence, the syndromes are functionally similar to LQT1.





**FIGURE 9.7** ECGs from three siblings with LQT2. Note that while the QT interval ranges from near normal ( $QT_c$  460 ms) to very prolonged ( $QT_c$  580 ms), the T wave morphology has similarities in all three. ECG, Electrocardiogram.

There seems to be a particular risk of arrhythmia during exercise and especially swimming. Exercise tolerance testing shows the absence of normal shortening of the QT interval on exercise and into the recovery phase. By contrast, in normal individuals the QT interval prolongs slightly at first but then rapidly shortens with increasing heart rate due to adrenaline-induced activation of  $i_{Ks}$ .  $\beta$ -blocker drugs are particularly effective in this condition and may obviate the need for an implantable cardioverter device (ICD) even in those who present with a cardiac arrest. The condition appears to be significantly less lethal than LQT2 or LQT3, in part no doubt due to the responsiveness to therapy but also the resting  $QT_c$  is more frequently normal [5]. The risk of cardiac events is higher in boys than girls, and this male preponderance of risk increases still further in adults [6].

#### 9.4.2.2.2 Long QT syndrome type 2

Mutations in *KCNH2* cause loss-of-function effects in the hERG (human ether-a-go-go related gene) channel, which carries the rapid delayed rectifying  $K^+$  current,  $i_{Kr}$ . The surface ECG tends to show slurred and notched low-amplitude T waves. The protein encoded by *KCNE2* (LQT6) is associated with the hERG channel. It usually does not cause clinical disease alone but is a potent cause of drug-induced long QT (discussed in Section 9.4.4).

Arrhythmias tend to be caused by sudden noise or emotional stress. The hERG channel is sensitive to both estrogen and progesterone, and some female patients report a change in symptoms depending upon their menstrual cycle. The postpartum state appears to be particularly risky, with an increase in sudden death in the 12 months following delivery. This risk appears to have been exacerbated by cessation of  $\beta$ -blockers, which was previously advised due to worries about intrauterine growth retardation. The risk of sudden death in both



genders appears to be equal until puberty then increases in women. Hypokalemia is dangerous and must be aggressively treated.

#### 9.4.2.2.3 Long QT syndrome type 3

LQTS type 3 is caused by gain-of-function mutations in *SCN5a*, which encodes the  $\alpha$ -subunit of the sodium-channel  $\text{Na}_v1.5$ . This increases  $i_{\text{Na}_v}$  leading to increased depolarizing current compared to repolarizing current, and hence prolongation of QT interval. The surface ECG shows a long flat ST segment, followed by a T wave that is relatively normal in morphology. The degree of QT prolongation in LQT3 has been considered less useful in terms of risk stratification than in LQT1 and 2 [5], though more recent evidence suggests that this is not the case [7]. The condition is rare compared to LQT1 or LQT2, and this increases the challenge of studying its clinical course. The first event tends to occur later in childhood but is more likely to be lethal; and hence, some authorities advocate ICD therapy even in asymptomatic patients with LQT3.

Mutations in *SCN5A* have also been implicated in BrS, progressive cardiac conduction disease, atrial arrhythmias, sinus node disease, and dilated cardiomyopathy. It is a particularly large gene with correspondingly substantial genetic noise, and so variants can be difficult to interpret. Patients with LQT3 are more likely to show inappropriate sinus bradycardia and flecainide will induce a Brugada ECG pattern in up to half of these patients. The proteins encoded by *CAV3* (LQT9), *SCN4B* (LQT10), and *SNTA1* (LQT12) interact with the  $\text{Na}_v1.5$  channel, and therefore these syndromes largely mimic LQT3.

#### 9.4.2.2.4 Jervell and Lange-Nielsen syndrome

This is an autosomal-recessive condition caused by homozygosity or compound heterozygosity in either *KCNQ1* or *KCNE1*, encoding the  $\text{K}_v7.1$   $\text{K}^+$  channel or its accessory protein minK1 (which otherwise cause LQT1 and LQT5, respectively). The syndrome is characterized by severe QT prolongation, frequent arrhythmia, and sensorineural deafness.

#### 9.4.2.2.5 Anderson–Tawil syndrome

Also termed LQT7, the majority of the cases of this syndrome have been found to be caused by a mutation in *KCNJ2*, which encodes  $\text{K}_{ir2.1}$  (an inward rectifier  $\text{K}^+$  channel). It is characterized by quite modest QT prolongation with a prominent U wave, short stature, characteristic facies, clinodactyly, and hypokalemic periodic paralysis.

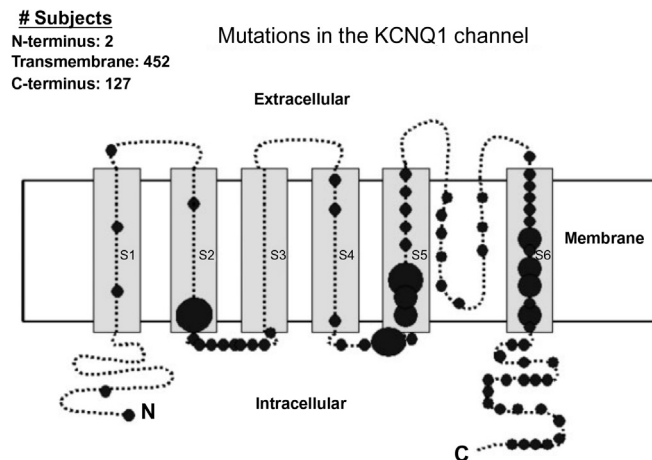
#### 9.4.2.2.6 Timothy syndrome

This syndrome is caused by mutations in *CACNA1C*, which encodes the  $\alpha$ -subunit of  $\text{Ca}_v1.2$  (an L-type  $\text{Ca}^{2+}$  channel). These patients display syndactyly, structural heart abnormalities (atrial and/or ventricular septal defects, patent foramen ovale or patent ductus arteriosus), and severe QT prolongation with frequent ventricular arrhythmias and an overall poor prognosis. Other arrhythmias such as atrial fibrillation, sinus bradycardia, and 2:1 heart block are also common. It is also termed LQT8.

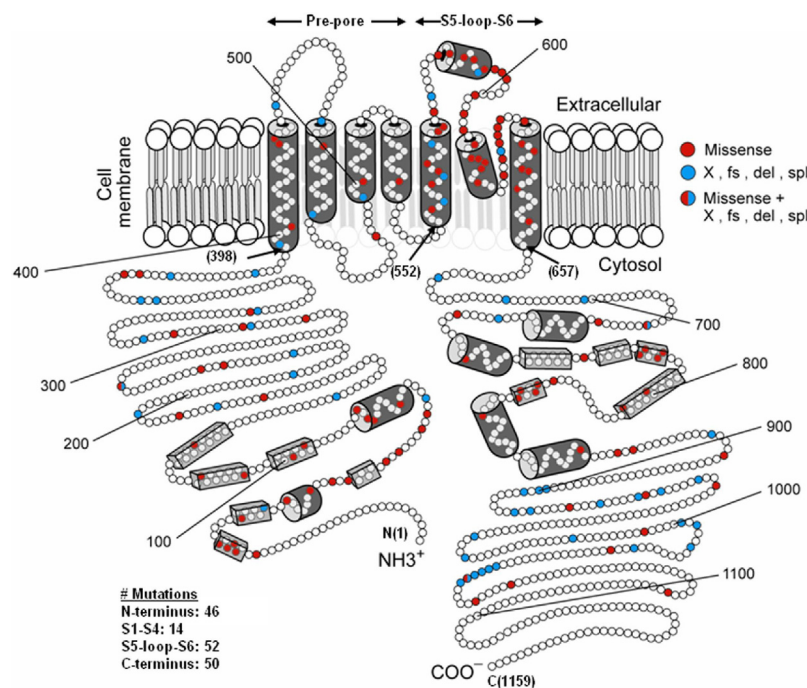
### 9.4.3 Molecular risk stratification in long QT syndrome

Moss et al. studied genotype–phenotype correlations in 600 patients with LQT1 [6]. Transmembrane pore-localizing mutations seem to carry higher risk than those in the C-terminus in LQT1. Such mutations may be predicted to disrupt the function of the channel to a greater degree. Mutations in the N-terminal region were rare, and very little can be said on their relative risk. Missense mutations had a higher rate of events than nonmissense mutations. Finally, dominant negative mutations reduce  $i_{\text{K}_s}$  by up to 84%, as compared to  $\leq 50\%$  for haploinsufficiency-type mutations, and carry a still higher risk in LQT1.

In a later study of 860 LQT1 patients (including some of the Rochester cohort), mutations in the cytoplasmic loops, which link the transmembrane regions of the pore, carried a particularly high risk in untreated patients but were associated in vitro with a strong response to  $\beta$ -blockade [8]. One particular pore-localizing mutation, A341V, appears to run an especially malignant clinical course. It is relatively common, and the reason for its malignancy is not well understood; it is described as having a mild dominant negative effect and only causes a modest reduction in  $i_{\text{K}_s}$  [9] (Fig. 9.8).



**FIGURE 9.8** Diagram of the KCNQ1 potassium channel, with LQT1-causing mutations marked. Position in the transmembrane pore region seems to carry higher risk than in the C-terminus. Circle size indicates number of affected individuals at each locus. Source: Reprinted with permission from Moss AJ, Shimizu W, Wilde AAM, Towbin JA, Zareba W, Robinson JL, et al. Clinical aspects of type-1 long-QT syndrome by location, coding type, and biophysical function of mutations involving the KCNQ1 gene. *Circulation* 2007;115(19):2481–9. <https://doi.org/10.1161/CIRCULATIONAHA.106.665406>.



**FIGURE 9.9** Diagram of the protein structure of the KCNH2 potassium channel, marked for known pathogenic mutations. Position in the transmembrane pore or N-terminus seems to carry higher risk. Source: Reprinted with permission from Shimizu W, Moss AJ, Wilde AAM, Towbin JA, Ackerman MJ, January CT, et al. Genotype-phenotype aspects of type 2 long QT syndrome. *J Am Coll Cardiol* 2009;54(22):2052–62. <https://doi.org/10.1016/j.jacc.2009.08.028>.

Common variants in the 3'-untranslated region of KCNQ1 have been found to affect the  $QT_c$  and to predict symptom occurrence in LQT1 [10]. This effect was allele specific, that is, the presence of these single nucleotide polymorphisms on the mutated allele was protective, but their presence was detrimental on the normal allele.

In a similar manner to that seen in LQT1, mutations in the transmembrane pore in LQT2 appear to carry a higher risk than those in the C-terminus [11]. In contrast to LQT1, more mutations were seen in the longer cytoplasmic N-terminal region, and these also appear to carry an increased risk (Fig. 9.9).

Given the substantially lower incidence of LQT3, it is difficult to draw conclusions about genotype–phenotype correlations. In the largest published cohort, two mutations in the C-terminal end of  $\text{Na}_v1.5$  (E1784K and D1790G) appeared to follow a relatively benign clinical course [7]. E1784K is known to confer an overlap phenotype with BrS, which is discussed in more detail later.

The Jervell and Lange-Neilsen syndrome is characterized by particularly frequent arrhythmia and a high risk of sudden death. In an analogous manner, “double-hit” LQTS—either from compound or digenic heterozygosity—also appear to have a worse prognosis [12]. Finally, a gene that encodes a nitric oxide synthase adaptor protein (*NOS1AP*) appears to act as a modifier; both affecting the corrected QT interval in patients with genetically determined LQTS and also, separately, predicting the risk of sudden death in those with a relatively normal  $\text{QT}_c$  ( $<500$  ms) [13]. These patients can be particularly hard to risk stratify by clinical means.

#### 9.4.4 Drug-induced long QT

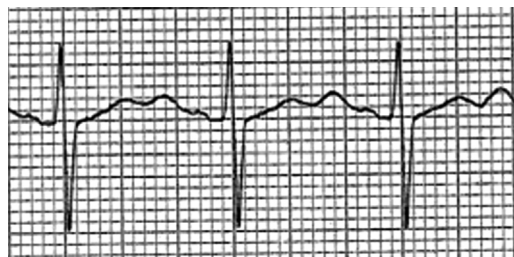
Any patient with diagnosed LQTS should avoid QT-prolonging drugs as they carry a risk of provoking arrhythmia. These are generally drugs that inhibit  $i_{K_r}$ , and new medications are tested for this property specifically. Lists of currently licensed drugs that carry risk are available for doctors and patients online (e.g., [www.crediblemeds.org](http://www.crediblemeds.org)). However, the most common cause of arrhythmia due to medications is in fact polygenic LQTS. An important concept is that of “repolarization reserve”; this is the innate resistance to QT-prolonging stimuli such as medications or electrolyte abnormalities and is variable between individuals. Certainly genetics play a major role in determining an individual’s risk of drug-induced arrhythmia, and work is ongoing to understand this serious public health risk in more detail. It seems likely that as part of a drive toward personalized medicine, genetic risk scoring to determine susceptibility will enter the clinical domain [14] (Fig. 9.10).

#### 9.4.5 Generation and propagation of arrhythmia in long QT syndrome

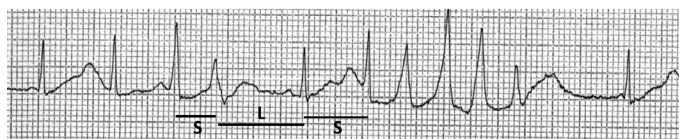
It is thought that arrhythmia generally begins with triggered activity caused by an early afterdepolarization. A re-entry circuit can then be set up due to unidirectional block in the middle layer of cardiomyocytes (M cells) and a differential refractoriness across the heart [15]. Under normal circumstances, there is an apicobasal repolarization gradient (predominantly in the epicardium) as well as an endo- to epicardial repolarization gradient. The M cells can prolong their action potential in response to rate slowing to a greater extent than epi- or endocardial layer. Conditions leading to reduction in  $i_{K_r}$  or increase in late  $i_{\text{Na}}$  tend to preferentially affect the M cell action potential duration. The dispersion of repolarization/refractoriness may be as important as the absolute QT duration [16], and this might explain the persistent vulnerability of some patients with LQTS despite a normal or nearly normal QT interval.

This process is clinically manifest as a ventricular ectopic, which lands in the vulnerable period (the “R-on-T” phenomenon), which leads to torsades-de-pointes. This most commonly occurs after a relative pause. The normal beat following a long R–R interval has a longer repolarization phase (comparable with the longer QT interval seen at lower heart rates), and an ectopic therefore has a higher risk of landing in the vulnerable period and causing arrhythmia. This progression of a normal interval, a longer interval (relative pause), and then an early ectopic is known as a “short–long–short” sequence (Fig. 9.11).

T wave alternans (T waves alternating either in height or between positive and negative on subsequent beats) is a relatively rare ECG feature but is felt to be a marker of increased risk. It may be present constantly or may



**FIGURE 9.10** A case of drug-induced long QT syndrome (provoked by methadone and alcohol withdrawal); the patient required temporary pacing for recurrent torsades-de-pointes with syncope. Note the similarity in morphology to LQT2, which is also caused by reduced  $i_{K_r}$ .



**FIGURE 9.11** ECG showing classic “short–long–short” sequence leading to 5 beats of non-sustained ventricular tachycardia. This is the typical pattern of arrhythmia initiation in Long QT syndrome. S, short; L, long.

**TABLE 9.3** Table summarizing the molecular basis and features of short QT syndrome.

	Gene	Current	Clinical features
SQT1	KCNH2	$\uparrow i_{Kr}$	All encode subunits of the L-type $Ca^{2+}$ channel. The ECG shows right praecordial ST elevation reminiscent of BrS
SQT2	KCNQ1	$\uparrow i_{Ks}$	
SQT3	KCNJ2	$\uparrow i_{K1}$	
SQT4	CACNA1C	$\downarrow i_{Ca}$	
SQT5	CANB2	$\downarrow i_{Ca}$	
SQT6	CACNA2D1	$\downarrow i_{Ca}$	

BrS, Brugada syndrome; ECG, electrocardiogram.

come and go. It is unclear exactly what causes this phenomenon, but it may be caused by early afterdepolarizations on alternating beats or maybe more likely by action potential duration restitution. This is a normal phenomenon—a long action potential duration leads to a long diastolic interval and thereby more active repolarizing channels in next cycle. T wave alternans may be seen prior to the generation of arrhythmia in other conditions (e.g., BrS and CPVT) prior or following an episode of ventricular arrhythmia. This association suggests a potential causative role. In addition, QT dispersion on the surface ECG (widely variable QT intervals in different leads) may reflect an increased or abnormal repolarization gradient across the heart and may also be a risk factor for arrhythmia.

## 9.5 Short QT syndrome

In the 1990s and early 2000s there was growing understanding of channelopathies as a group of genetic diseases, which stimulated clinical research. Early reports of patients with abnormally short QT intervals and an association with AF and with sudden cardiac death led to the description of a new syndrome [17], and a genetic basis was rapidly determined [18,19]. The ECG generally shows tall, peaked T waves without a flat ST segment and a  $QT_c \leq 360$  ms in men and  $\leq 370$  ms in women (although the  $QT_c$  is usually  $\leq 320$  ms in all). It is inherited in an autosomal-dominant manner, and all forms are very rare—only around 200 patients have been described in the literature so far. The natural history is consequently not well known, with a risk of sudden cardiac death of approximately 1% per annum. Most episodes occur during sleep or at rest.

The yield of genetic testing is lower than in LQTS at between 15% and 40%. When a genotype is found, the syndrome is most commonly caused by gain-of-function mutations in  $K^+$  channels that carry repolarizing current (SQT1-3), or loss-of-function mutations in  $Ca^{2+}$  channels leading to a reduction in depolarizing currents in the plateau phase (SQT4-6) (Table 9.3).

### 9.5.1 Generation and propagation of arrhythmia

The action potential and refractory period of cardiomyocytes is reduced in SQTS, which increases the vulnerable period, especially at low heart rates. In addition, increased early repolarization current may cause loss of the action potential dome and phase 2 re-entry in a manner analogous to that proposed for Brugada syndrome.

## 9.6 Catecholaminergic polymorphic ventricular tachycardia

CPVT is described as ventricular tachyarrhythmia (typically bidirectional VT) in conditions of exercise or emotional stress in a structurally normal heart. The resting ECG is normal, and ischemic heart disease should be excluded. It is a rare disease, and the estimated prevalence is approximately 1:10,000. The onset of symptoms most frequently occurs in children and young adults, and there is often a family history of recurrent syncope and/or sudden death.

CPVT is dominantly inherited, and the yield of genetic testing is moderately high, with a pathogenic variant discovered in approximately 60% of those tested. In 1999 Swan et al. mapped the locus to chromosome 1q42-q43, which is the position of *RYR2*, the gene encoding the ryanodine receptor (RyR2) [20]. RyR2 is the calcium channel in the membrane of the SR, which mediates  $\text{Ca}^{2+}$ -induced  $\text{Ca}^{2+}$  release. Several mutations affecting *RYR2* were subsequently described [21], as well as the discovery of an autosomal-recessive form, linked to missense mutations in *CASQ2*, which encodes calsequestrin [22]. Calsequestrin is the major  $\text{Ca}^{2+}$  storage protein in the SR and also acts as a luminal  $\text{Ca}^{2+}$  sensor, preventing RyR2 from opening when stores of  $\text{Ca}^{2+}$  are low. An autosomal-dominant form of CPVT associated with a mutation in *CASQ2* has recently been described [23]. Calsequestrin links to RyR2 via the transmembrane junctin–triadin complex. Mutations in *TRDN*, which encodes triadin, have also been linked to a recessive form of CPVT [24].

RyR2 has a large cap, which extends into the cytosol and interacts with a further group of regulatory proteins, including calmodulin (CaM). CaM is essential and is ubiquitously expressed; it has structural similarities to troponin c, another  $\text{Ca}^{2+}$ -binding molecule. CaM interacts with a wide variety of different pathways, but its role in cardiomyocytes seems to be inhibiting  $\text{Ca}^{2+}$  release from the SR in response to high cytosolic  $\text{Ca}^{2+}$ . Severe forms of CPVT and LQTS causing recurrent cardiac arrest and death in infancy have been linked to missense mutations in *CALM1*, *CALM2*, or recently *CALM3*, three genes that each encode CaM [25–27]. These patients may have features of either LQTS or CPVT or both, and they do seem to respond to  $\beta$ -blockers. Less severe forms of disease are increasingly recognized, and downregulation of CaM is thought to be an important mechanism of arrhythmogenesis in other conditions such as heart failure [28].

### 9.6.1 Generation and propagation of arrhythmia in catecholaminergic polymorphic ventricular tachycardia

The central mechanism of arrhythmogenesis in CPVT appears to be defective control of cytosolic calcium, caused most commonly by leakage from the SR stores through RyR2 but also by failure of downstream control.  $\beta$ -adrenergic stimulation by stress or exercise increases  $\text{Ca}^{2+}$  flux through L-type calcium channels and increases  $\text{Ca}^{2+}$  concentration in the SR. This causes opening of RyR2 via stimulation of calsequestrin, through the junctin–triadin complex. It can also cause RyR2 opening directly by the release of various molecular brakes. CPVT-causing mutations tend to favor RyR2 opening at lower  $\text{Ca}^{2+}$  concentration, cause leak throughout systole and diastole, or release the channel from control by calsequestrin or calmodulin. Local cytosolic spikes in  $\text{Ca}^{2+}$ , called “calcium sparks,” cause activation of the  $\text{Na}^+$ – $\text{Ca}^{2+}$  exchange pump (NCX) and an increase in  $i_{\text{Na}}$ . Activation of NCX classically also causes bidirectional ventricular tachycardia (VT) in digoxin toxicity.

Increased  $i_{\text{Na}}$  leads to delayed (phase 4) afterdepolarizations, which become visible as ectopics on the surface ECG. These ectopics predominantly arise in the epicardium, which sets up a repolarization gradient and therefore a substrate for re-entry and polymorphic VT [29]. Alternatively, separate right ventricle (RV) and left ventricle (LV) foci have been proposed as the basis of bidirectional VT. These foci are at the insertion points of the major Purkinje system branches, which are particularly prone to delayed afterdepolarizations and ectopy [2].

### 9.6.2 Molecular risk stratification in catecholaminergic polymorphic ventricular tachycardia

There are three isoforms of ryanodine receptor in the human (RyR1-3); they are each a homotetramer (constructed of four protomers) and appear to be shaped like a mushroom, with a stalk crossing the SR membrane and a large (~80% by mass) cap bulging into the cytosol [30]. There is a particular hotspot for mutations in an anion-binding site in the N-terminal region, which appears under normal circumstances to act as a brake for channel opening [31]. However, mutations in channel-forming region may carry more arrhythmia risk than those in the N-terminal. Recessive forms of CPVT linked to mutations in *CASQ2* or *TRDN* are relatively rare, and



little is known about genotype–phenotype correlation. However, CPVT caused by mutations in CALM1 or CALM2 are frequently severe, causing recurrent cardiac arrest and death in infancy.

## 9.7 Brugada syndrome

The association of ventricular arrhythmia with typical ECG changes was initially described in 1989 [32], and the condition was further elucidated by and eventually named after Josep Brugada and Pedro Brugada [33]. BrS is described as polymorphic VT/ventricular fibrillation (VF) in the context of right precordial coved ST elevation, negative T waves, and RV delay (the type 1 Brugada pattern). The peak incidence of ventricular arrhythmia is in the fourth and fifth decades of life and is far more common in males (preponderance of 3:1 worldwide, and up to 9:1 in Thailand). Arrhythmia tends to occur at rest or during sleep, in contrast to most other arrhythmia syndromes.

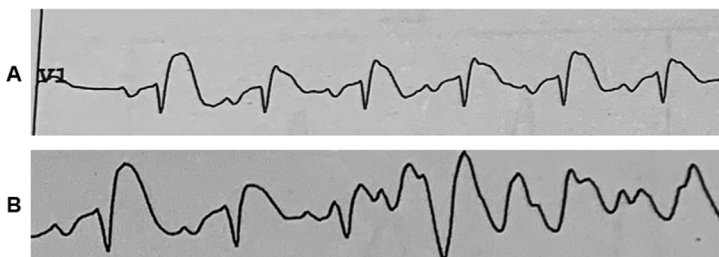
The characteristic (“type 1”) ECG pattern may come and go during the day and may be present at rest or only seen after provocation with sodium channel–blocking drugs (e.g., ajmaline, procainamide and flecainide). Fever or excessive alcohol may also bring out these ECG changes and increase the risk of arrhythmia. An online list of drugs that may increase risk in this condition may be found on the website: [www.brugadadrugs.org](http://www.brugadadrugs.org), and patients and their health professionals should be encouraged to avoid these. The risk of sudden cardiac death appears to be higher in those with prior cardiac arrest, a spontaneous (rather than drug-induced) type 1 pattern on the ECG, and in those with prior syncope. However, numerically, most cardiac arrest events are probably in the “low-risk” population so current risk stratification requires improvement [34].

The yield of genetic testing is low, with a pathogenic mutation found in approximately 20% of individuals tested. Described mutations cause loss of sodium-channel function by changing the opening and closing kinetics or haploinsufficiency and abnormal trafficking.

The picture is further complicated by the observation that clinically manifest disease does not always segregate with pathogenic mutations in affected families. Variable penetrance may be the explanation in certain individuals, but clinically affected relatives who lack the familial mutation have also been reported. This is more difficult to explain. Recent work suggests that common variants affect disease expression [35] and that BrS might in fact be considered an oligogenic disease, with expression strongly linked to age and other environmental factors. The definition of the disease has changed over the years as diagnostic criteria have developed. In the latest criteria, only symptomatic individuals (or those with a family history of the condition or sudden death) with concealed or manifest ECG changes are given a formal diagnosis, but previous criteria allowed a diagnosis on drug provocation testing alone. It is not clear how many of the population will show type 1 ECG changes in response to drug provocation (a “positive ajmaline test”, Fig. 9.12); occasionally people with a negative test will become positive on re-challenge, and the rate of positive responses in the normal population may be substantially higher than previously recognized. A study in patients undergoing ablation for an atrial arrhythmia (atrioventricular nodal re-entry tachycardia, AVNRT) showed a positive test in 27% of arrhythmia patients and 4.5% of controls [36]. Even given a possible link between BrS and AVNRT, the high rate in control subjects suggests that a positive response to drug challenge (in the absence of other features of BrS) may be relatively common in the normal population and may not confer a high risk of arrhythmia (Fig. 9.12).

### 9.7.1 Generation and propagation of arrhythmia in Brugada syndrome

There are two prevailing theories of arrhythmia generation in BrS, termed the “depolarization” and “repolarization” theories. It is important to recognize that the two theories are not mutually exclusive and may in fact be



**FIGURE 9.12** ECG showing ajmaline challenge test for BrS. The test was positive showing a type 1 ECG pattern in lead V1 (A) with T wave alternans suggesting electrical instability. The patient subsequently developed ventricular fibrillation (B) and was successfully resuscitated.



complementary, they are reviewed in [37]. The depolarization theory points to conduction delay within the right ventricular outflow tract (RVOT) compared the remainder of the myocardium. A ventricular ectopic beat sets up a macro re-entry circuit and then VF. Evidence for this theory comes largely from electrophysiological studies in patients. The repolarization theory postulates that the deep action potential notch (otherwise known as “loss of the action potential dome”) due to decreased early  $i_{Na}$  or increased  $i_{to}$  leads to a phenomenon called “phase 2 re-entry.” Essentially, the relative prominence of the action potential (AP) notch in the epicardium sets up a severe transmural repolarization gradient and allows re-entry between the different layers of myocardium; it does not require macroscopic circus movement. This is visible on the surface ECG as a short-coupled ectopic, which leads to VF/polymorphic VT. Evidence for this theory largely comes from animal models such as the canine myocardial wedge model. Aging leads to fibrosis and reduced gap junction connections [38], this may explain the increase in arrhythmias in the fourth/fifth decades and would favor the depolarization hypothesis. This type of progressive fibrosis is also thought to be important in the generation of arrhythmias in cardiomyopathies (such as dilated and arrhythmogenic cardiomyopathies).

## 9.8 Targeted molecular diagnosis and therapy

Better understanding of the molecular basis and pathophysiology of inherited arrhythmia syndromes has led to the adoption of targeted therapies (lifestyle, pharmacological, and even surgical). One of the most important decisions in the management of patients with inherited arrhythmia syndromes is whether or not to implant an ICD. ICDs can save lives in a variety of conditions but carry a substantial risk of complications including inappropriate shocks, psychological problems and depression, restrictions to lifestyle and driving and device-related complications such as generator changes, or infection/lead fracture requiring extraction (and this procedure carries a 1% risk of mortality). Risk stratification in the individual patient is therefore essential. At present, this relies predominantly on clinical factors, though knowledge of molecular pathology and the behavior of specific genetic lesions and modifiers will be of increasing importance in the future.

Genetic testing by Sanger sequencing is slow and expensive, and particularly challenging for certain large genes that are relevant to inherited cardiac conditions, such as *RYR2* and *SCN5A*. The mainstay of genetic diagnosis in many centers is now next-generation sequencing (NGS), testing panels of genes (i.e., cardiomyopathy or arrhythmia panels), or whole-exome testing with reporting of large numbers of genes associated with heart disease. Whole-genome sequencing may be done in specific challenging patients and families, or as part of research/clinical projects such as Genetics England’s “100,000 Genomes Project”. At present, testing is limited by cost, and there is a balance between read depth, which is best with NGS, and coverage of the genome. The majority of cardiac disease-causing variants are private and so interpreting variants of uncertain significance is a major challenge.

### 9.8.1 Long QT syndrome

#### 9.8.1.1 Clinical risk assessment in long QT syndrome

There is increasing understanding of genotype–phenotype correlations in LQTS but expression of the disease in a given individual varies widely. Undoubtedly, there are undiscovered genetic factors underlying this variable expression but for now the risk stratification depends primarily upon clinical features. Corrected QT interval ( $QT_c$ ) correlates with risk of arrhythmia in LQT1-3, and expert consensus recommends that a  $QT_c > 550$  ms on therapy is a reasonable threshold for the discussion of primary prevention ICD implantation in adults. Evidence in rarer syndromes is lacking but LQT8 (Timothy syndrome) and Jervell and Lange-Neilsen syndrome seem to carry an exceptionally high risk of arrhythmia. Syncope or cardiac arrest, especially on  $\beta$ -blocker therapy, predicts future episodes and is an indication for implantation of an ICD. Family history is more difficult to interpret. Given the large difference in disease expression, there is little evidence that severe disease in a family member predicts episodes in an affected relative. However, discussions about risks and benefits of ICD therapy will clearly be affected by the family experience of disease.

#### 9.8.1.2 Targeted therapies in long QT syndrome

All confirmed cases of LQTS should be advised to avoid QT-prolonging drugs (listed on the website [www.crediblemeds.com](http://www.crediblemeds.com) and on their smartphone app) and given tailored exercise advice. The first-line

pharmacological treatment is with a noncardioselective  $\beta$ -blocker drug such as nadolol, propranolol, or carvedilol. This is advisable in all those with confirmed LQTS (either clinically or genetically) and even in genotype-positive–phenotype-negative patients in the absence of a prolonged QT interval. Early studies in LQT3 failed to show a benefit of  $\beta$ -blockers, and the high incidence of bradycardia and heart block (along with evidence from in vitro studies) raised the possibility that they might be proarrhythmic or cause death by bradyarrhythmia. This has been investigated in a large cohort, and  $\beta$ -blockers do appear to be safe and effective in LQT3 [7].

LQTS3 is caused by gain-of-function mutations in SCN5A, with increased  $i_{Na}$ . Therefore sodium-channel blockers such as mexiletine and flecainide might be postulated to be effective in this condition. On challenge with flecainide up to 50% of patients with LQT3 develop a Brugada type 1 pattern on ECG so caution should be exercised [39]. Mexiletine seems to be less likely than flecainide to induce ECG changes, this may be because flecainide blocks both the early peak  $i_{Na}$  as well as the late  $i_{Na}$  and also has effects on  $i_{Kr}$ . There is now evidence of safety and efficacy of mexiletine in a cohort of 34 LQT3 patients [40]. Finally, ranolazine has been suggested as a potentially beneficial therapy in this condition. It preferentially blocks late  $i_{Na}$  with substantially less effect on early peak  $i_{Na}$ , so theoretically at least should be less likely to induce a Brugada pattern on the ECG. There is evidence from experimental models that it reduced late  $i_{Na}$  and shortened the  $QT_c$  in eight patients with LQT3 without inducing a Brugada ECG pattern [41].

Adrenaline and the sympathetic nervous system appear to be particularly important in arrhythmogenesis in LQT1, as evidenced by the increase in episodes of syncope and sudden death on exercise. Swimming and the act of jumping into cold water are especially potent activators of the sympathetic nervous system. The heart receives its sympathetic innervation from the left stellate ganglion and the four or five upper thoracic ganglia on the left. The LCSD procedure has developed over time from removal of the stellate ganglion alone (which causes Horner's syndrome and does not completely denervate the heart) to additionally removing four or five of the upper thoracic ganglia. Only the lower half of the stellate ganglion actually needs to be removed, and this decreases the incidence of postoperative Horner's syndrome [42]. Infiltration of the ganglia with local anesthetic at the beginning of the operation reduces the risk of arrhythmic storm due to sympathetic activation. LCSD has been shown to be effective in LQT1, and its use has been extended to other variants of LQTS, albeit with slightly a lower predicted rate of success.

## 9.8.2 Short QT syndrome

### 9.8.2.1 Clinical risk assessment and therapy in short QT syndrome

SQTS appears to be highly penetrant with an annual incidence of sudden cardiac death of approximately 1%. Implantation of an ICD should be considered in all survivors of cardiac arrest and is recommended in confirmed cases with documented sustained ventricular arrhythmia, even in the absence of syncope. Expert consensus suggests quinidine or sotalol may be useful in asymptomatic patients with a family history of sudden death or to prevent therapies as an adjunct to an ICD [43]. Ranolazine, with its effects on late  $i_{Na}$  and  $i_{Kr}$ , might prolong repolarization and reduce arrhythmia and there is evidence from animal models that this might be the case [44].

## 9.8.3 Catecholaminergic polymorphic ventricular tachycardia

### 9.8.3.1 Clinical risk assessment and therapy in catecholaminergic polymorphic ventricular tachycardia

All forms of CPVT carry a high risk of arrhythmia and sudden death; around 40% of patients die within 10 years of diagnosis [45]. The critical role of the sympathetic nervous system in arrhythmogenesis in CPVT has led to a recommendation that any confirmed case should be given the maximum tolerated dose of  $\beta$ -blockers. There is evidence that  $\beta$ -blockade reduces symptoms and sudden death (with five times more events in untreated patients).  $\beta$ -blockers reduce  $Ca^{2+}$  accumulation in the SR, thereby decreasing RyR2 opening and  $Ca^{2+}$  release. A further targeted therapy is LCSD, which nearly abolished symptoms in a group of CPVT patients who were on average receiving >3 ICD shocks per annum on maximal medical therapy [46].

Some authorities recommend avoiding ICDs as long as possible, as there have been reports of ventricular electrical storm being provoked by shocks (and hence massively increased sympathetic drive), and ICDs are by no means always successful [47]. On the other hand, even on maximal  $\beta$ -blockade the rate of sudden death and disability is substantial. More recently, sodium-channel blockers (flecainide and mexiletine) have been found to be effective in CPVT. Sodium-channel blockers work in a complementary manner downstream of the action of  $\beta$ -blockers. They reduce the likelihood of an action potential resulting from the early afterdepolarizations seen in

CPVT (damping down the excitability of the cell, known as negative bathmotropy) and may have a direct effect on the RyR2 channel. There is some evidence that flecainide monotherapy is beneficial in those who cannot tolerate  $\beta$ -blockers (most commonly due to asthma) [48]. Finally, bidirectional VT has been suggested to arise in the Purkinje fibers (as discussed earlier), and successful catheter ablation at these sites has been reported [49].

### 9.8.4 Brugada syndrome

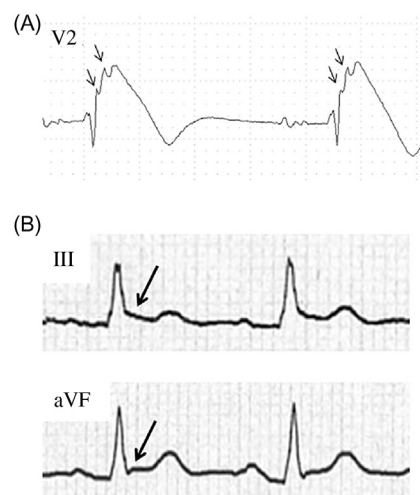
#### 9.8.4.1 Clinical risk assessment in Brugada syndrome

Patients with a spontaneous type 1 pattern on the surface ECG carry more risk than those with a latent type 1 pattern. The ECG may change through the day, and 12-lead Holter monitoring can be useful in increasing sensitivity for this. Prior syncope and a family history of sudden death in a first-degree relative are considered risk factors. On the 12-lead ECG, QRS fragmentation and early repolarization seem to carry an increased risk. QRS fragmentation in the right precordial leads may reflect a greater degree of depolarization delay in the RVOT, whereas early repolarization in the inferior leads is likely to reflect a more widespread abnormality of repolarization [50] (Fig. 9.13).

#### 9.8.4.2 Targeted therapy in Brugada syndrome

Along with the avoidance of dangerous drugs ([www.brugadadrugs.org](http://www.brugadadrugs.org)), aggressive management of fever, and avoidance of excessive alcohol, management can involve the implantation of a defibrillator. This is clearly indicated in survivors of cardiac arrest (secondary prevention), but primary prevention is a controversial area with an increasing recognition that those with a type 1 pattern on drug challenge only, and no other risk factors, are probably at little more than background risk of sudden death. Quinidine has a role in some patients; as an  $i_{to}$  blocker, it normalizes the AP notch and can therefore prevent phase 2 re-entry. It is usually used as an adjunct in patients who have required shocks from their ICD, in an effort to prevent further painful shocks.

Recently, there have been reports of the use of electrophysiological studies and ablation in nine patients with recurrent VF and ICD shocks (averaging 4 per month). The endocardial and epicardial surfaces of the RVOT were mapped. Signals from the endocardial surface were normal, but those from the epicardial surface showed low voltage (<1 mV), prolonged duration, and late polyphasic potentials. Ablation targeting these signals normalized the ECG and prevented ventricular arrhythmia [51]. Such improvement would be consistent with either the depolarization or repolarization hypotheses.



**FIGURE 9.13** Examples of (A) QRS fragmentation and (B) inferior early repolarization in Brugada syndrome. These ECG features are relatively uncommon but, when present, are thought to confer increased risk. ECG, Electrocardiogram. Source: Reprinted with permission from Tokioka K, Kusano KF, Morita H, Miura D, Nishii N, Nagase S, et al. Electrocardiographic parameters and fatal arrhythmic events in patients with brugada syndrome: combination of depolarization and repolarization abnormalities. *J Am Coll Cardiol* 2014;63(20):2131–8. <https://doi.org/10.1016/j.jacc.2014.01.072>.

## 9.9 Summary

Increasing knowledge of genotype–phenotype correlations and important modifiers (particularly in LQTS and CPVT) may assist in risk stratification, but for now individual clinical assessment is the mainstay. Understanding of the molecular pathology of these conditions has led to some specific therapies (e.g., mexiletine for LQT3 and quinidine for BrS), but we are a long way from routine use of targeted therapies in the majority of patients.

## References

- [1] Morita H, Wu J, Zipes DP. The QT syndromes: long and short. *Lancet* 2008;372(9640):750–63. Available from: [https://doi.org/10.1016/S0140-6736\(08\)61307-0](https://doi.org/10.1016/S0140-6736(08)61307-0).
- [2] Watanabe H, Knollmann BC. Mechanism underlying catecholaminergic polymorphic ventricular tachycardia and approaches to therapy. *J Electrocardiol* 2011;44(6):650–5. Available from: <https://doi.org/10.1016/j.jelectrocard.2011.07.025>.
- [3] Schwartz PJ, Crotti L. QTc behavior during exercise and genetic testing for the long-QT syndrome. *Circulation*. 2011;124:2181–4.
- [4] Barc J, Bricc F, Schmitt S, Kyndt F, Le Cunff M, Baron E, et al. Screening for copy number variation in genes associated with the long QT syndrome: clinical relevance. *J Am Coll Cardiol* 2011;57(1):40–7. Available from: <https://doi.org/10.1016/j.jacc.2010.08.621>.
- [5] Napolitano C, Ph D, Bloise R, Ronchetti E, Grillo M, Vicentini A, et al. Risk stratification in the long-QT syndrome. *N Engl J Med* 2003;348:1866–74.
- [6] Moss AJ, Shimizu W, Wilde AAM, Towbin JA, Zareba W, Robinson JL, et al. Clinical aspects of type-1 long-QT syndrome by location, coding type, and biophysical function of mutations involving the KCNQ1 gene. *Circulation* 2007;115(19):2481–9. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.106.665406>.
- [7] Wilde AA, Moss AJ, Kaufman ES, Shimizu W, Peterson DR, Benhorin J, et al. Clinical aspects of type 3 long-QT syndrome: an international multicenter study. *Circulation* 2016;134(12):872–82. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.116.021823>.
- [8] Wen B, Lampe JN, Roberts AG, Atkins WM, Rodrigues AD, Nelson SD. Mutations in cytoplasmic loops of the KCNQ1 channel and the risk of life-threatening events: implications for mutation-specific response to beta-blocker therapy in type-1 long QT syndrome. *Circulation* 2012;125:1988–96. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.111.048041>.
- [9] Crotti L, Spazzolini C, Schwartz PJ, Shimizu W, Denjoy I, Schulze-Bahr E, et al. The common long-QT syndrome mutation KCNQ1/A341V causes unusually severe clinical manifestations in patients with different ethnic backgrounds: toward a mutation-specific risk stratification. *Circulation* 2007;116(21):2366–75. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.107.726950>.
- [10] Amin AS, Giudicessi JR, Tijssen AJ, Spanjaart AM, Reckman YJ, Klemens CA, et al. Variants in the 3′ untranslated region of the KCNQ1-encoded Kv7.1 potassium channel modify disease severity in patients with type 1 long QT syndrome in an allele-specific manner. *Eur Heart J* 2012;33(6):714–23. Available from: <https://doi.org/10.1093/eurheartj/ehr473>.
- [11] Shimizu W, Moss AJ, Wilde AAM, Towbin JA, Ackerman MJ, January CT, et al. Genotype-phenotype aspects of type 2 long QT syndrome. *J Am Coll Cardiol* 2009;54(22):2052–62. Available from: <https://doi.org/10.1016/j.jacc.2009.08.028>.
- [12] Wilde AAM. Long QT syndrome: a double hit hurts more. *Heart Rhythm* 2010;7(10):1419–20. Available from: <https://doi.org/10.1016/j.hrthm.2010.06.027>.
- [13] George AL, Crotti L, Monti MC, Insolia R, Peljto A, Goosen A, et al. NOS1AP is a genetic modifier of the long-QT syndrome. *Circulation* 2009;120(17):1657–63. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.109.879643>.
- [14] Strauss DG, Vicente J, Johannesen L, Blinova K, Mason JW, Weeke P, et al. A common genetic variant risk score is associated with drug-induced QT prolongation and torsade de pointes risk: a pilot study. *Circulation*. 2017;116:023980. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.116.023980>.
- [15] Akar FG, Yan GX, Antzelevitch C, Rosenbaum DS. Unique topographical distribution of M cells underlies reentrant mechanism of torsade de pointes in the long-QT syndrome. *Circulation* 2002;105(10):1247–53. Available from: <https://doi.org/10.1161/hc1002.105231>.
- [16] Salama G, Akar FG. Deciphering arrhythmia mechanisms: tools of the trade. *Card Electrophysiol Clin* 2011. Available from: <https://doi.org/10.1016/j.ccep.2010.10.013>.
- [17] Gaita F, Giustetto C, Bianchi F, Wolpert C, Schimpf R, Riccardi R, et al. Short QT syndrome: a familial cause of sudden death. *Circulation* 2003;108(8):965–70. Available from: <https://doi.org/10.1161/01.CIR.0000085071.28695.C4>.
- [18] Brugada R, Hong K, Dumaine R, Cordeiro J, Gaita F, Borggrefe M, et al. Sudden death associated with short-QT syndrome linked to mutations in HERG. *Circulation* 2004;109(1):30–5. Available from: <https://doi.org/10.1161/01.CIR.0000109482.92774.3A>.
- [19] Hong K, Bjerregaard P, Gussak I, Brugada R. Short QT syndrome and atrial fibrillation caused by mutation in KCNH2. *J Cardiovasc Electrophysiol* 2005;16(4):394–6. Available from: <https://doi.org/10.1046/j.1540-8167.2005.40621.x>.
- [20] Swan H, Piippo K, Viitasalo M, Heikkilä P, Paavonen T, Kainulainen K, et al. Arrhythmic disorder mapped to chromosome 1q42-q43 causes malignant polymorphic ventricular tachycardia in structurally normal hearts. *J Am Coll Cardiol* 1999;34(7):2035–42. Available from: [https://doi.org/10.1016/S0735-1097\(99\)00461-1](https://doi.org/10.1016/S0735-1097(99)00461-1).
- [21] Priori SG, Napolitano C, Tiso N, Memmi M, Vignati G, Bloise R, et al. Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. *Circulation* 2001;103(2):196–200. Available from: <https://doi.org/10.1161/01.CIR.103.2.196>.
- [22] Lahat H, Pras E, Eldar M. A missense mutation in CASQ2 is associated with autosomal recessive catecholamine-induced polymorphic ventricular tachycardia in Bedouin families from Israel. *Ann Med* 2004;36(Suppl 1):87–91. Available from: <https://doi.org/10.1080/17431380410032517>.
- [23] Gray B, Bagnall RD, Lam L, Ingles J, Turner C, Haan E, et al. A novel heterozygous mutation in cardiac calsequestrin causes autosomal dominant catecholaminergic polymorphic ventricular tachycardia. *Heart Rhythm* 2016;13(8):1652–60. Available from: <https://doi.org/10.1016/j.hrthm.2016.05.004>.



- [24] Roux-buisson N, Cacheux M, Fourest-lieuvin A, Fauconnier J, Brocard J, Denjoy I, et al. Absence of triadin, a protein of the calcium release complex, is responsible for cardiac arrhythmia with sudden death in human. *Hum Mol Genet* 2012;21(12):2759–67. Available from: <https://doi.org/10.1093/hmg/dds104>.
- [25] Makita N, Yagihara N, Crotti L, Johnson CN, Beckmann BM, Roh MS, et al. Novel calmodulin mutations associated with congenital arrhythmia susceptibility. *Circ Cardiovasc Genet* 2014;7(4):466–74. Available from: <https://doi.org/10.1161/CIRCGENETICS.113.000459>.
- [26] Nyegaard M, Overgaard MT, Sondergaard MT, Vranas M, Behr ER, Hildebrandt LL, et al. Mutations in calmodulin cause ventricular tachycardia and sudden cardiac death. *Am J Hum Genet* 2012;91(4):703–12. Available from: <https://doi.org/10.1016/j.ajhg.2012.08.015>.
- [27] Crotti L, Johnson CN, Graf E, De Ferrari GM, Cuneo BF, Ovidia M, et al. Calmodulin mutations associated with recurrent cardiac arrest in infants. *Circulation* 2013;127(9):1009–17. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.112.001216>.
- [28] Yang Y, Guo T, Oda T, Chakraborty A, Chen L, Uchinoumi H, et al. Cardiac myocyte Z-line calmodulin is mainly RyR2-bound, and reduction is arrhythmogenic and occurs in heart failure. *Circ Res* 2014;114(2):295–306. Available from: <https://doi.org/10.1161/CIRCRESAHA.114.302857>.
- [29] Nam GB, Burashnikov A, Antzelevitch C. Cellular mechanisms underlying the development of catecholaminergic ventricular tachycardia. *Circulation* 2005;111(21):2727–33. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.104.479295>.
- [30] Zalk R, Clarke OB, des Georges A, Grassucci RA, Reiken S, Mancina F, et al. Structure of a mammalian ryanodine receptor. *Nature* 2015;517(7532):44–9. Available from: <https://doi.org/10.1038/nature13950>.
- [31] Kimlicka L, Tung C-C, Carlsson A-CC, Lobo PA, Yuchi Z, Van Petegem F. The cardiac ryanodine receptor N-terminal region contains an anion binding site that is targeted by disease mutations. *Structure* 2013;21(8):1440–9. Available from: <https://doi.org/10.1016/j.str.2013.06.012>.
- [32] Martini B, Nava A, Thiene G, Buja GF, Canciani B, Scognamiglio R, et al. Ventricular fibrillation without apparent heart disease: description of six cases. *Am Heart J* 1989;118(6):1203–9. Available from: [https://doi.org/10.1016/0002-8703\(89\)90011-2](https://doi.org/10.1016/0002-8703(89)90011-2).
- [33] Brugada P, Brugada J. Right bundle branch block, persistent ST segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome. A multicenter report. *J Am Coll Cardiol* 1992;20(6):1391–6. Available from: [https://doi.org/10.1016/0735-1097\(92\)90253-J](https://doi.org/10.1016/0735-1097(92)90253-J).
- [34] Raju H, Papadakis M, Govindan M, Bastiaenen R, Chandra N, O'Sullivan A, et al. Low prevalence of risk markers in cases of sudden death due to Brugada syndrome: relevance to risk stratification in Brugada syndrome. *J Am Coll Cardiol* 2011;57(23):2340–5. Available from: <https://doi.org/10.1016/j.jacc.2010.11.067>.
- [35] Bezzina CR, Barc J, Mizusawa Y, Remme CA, Gourraud J-B, Simonet F, et al. Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death. *Nat Genet* 2013;45(9):1044–9. Available from: <https://doi.org/10.1038/ng.2712>.
- [36] Hasdemir C, Payzin S, Kocabas U, Sahin H, Yildirim N, Alp A, et al. High prevalence of concealed Brugada syndrome in patients with atrioventricular nodal reentrant tachycardia. *Heart Rhythm* 2015;12(7):1584–94. Available from: <https://doi.org/10.1016/j.hrthm.2015.03.015>.
- [37] Sieira J, Dendramis G, Brugada P. Pathogenesis and management of Brugada syndrome. *Nat Rev Cardiol* 2016;13(12):744–56. Available from: <https://doi.org/10.1038/nrcardio.2016.143>.
- [38] Van Veen TAB, Stein M, Royer A, Le Quang K, Charpentier F, Colledge WH, et al. Impaired impulse propagation in Scn5a-knockout mice: combined contribution of excitability, connexin expression, and tissue architecture in relation to aging. *Circulation* 2005;112(13):1927–35. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.105.539072>.
- [39] Priori SG, Napolitano C, Schwartz PJ, Bloise R, Crotti L, Ronchetti E. The elusive link between LQT3 and Brugada syndrome: the role of flecainide challenge. *Circulation* 2000;102(9):945–7. Available from: <https://doi.org/10.1161/01.CIR.102.9.945>.
- [40] Mazzanti A, Maragna R, Faragli A, Monteforte N, Bloise R, Memmi M, et al. Gene-specific therapy with mexiletine reduces arrhythmic events in patients with long QT syndrome type 3. *J Am Coll Cardiol* 2016;67(9):1053–8. Available from: <https://doi.org/10.1016/j.jacc.2015.12.033>.
- [41] Chorin E, Hu D, Antzelevitch C, Hochstadt A, Belardinelli L, Zeltser D, et al. Ranolazine for congenital long-QT syndrome type III: experimental and long-term clinical data. *Circ Arrhythmia Electrophysiol* 2016;9(10):e004370. Available from: <https://doi.org/10.1161/CIRCEP.116.004370>.
- [42] Cho Y. Left cardiac sympathetic denervation: an important treatment option for patients with hereditary ventricular arrhythmias. *J Arrhythmia* 2016;32(5):340–3. Available from: <https://doi.org/10.1016/j.joa.2015.08.002>.
- [43] Hammond-Haley M, Patel RS, Providência R, Lambiase PD. Exercise restrictions for patients with inherited cardiac conditions: current guidelines, challenges and limitations. *Int J Cardiol* 2016;209:234–41. Available from: <https://doi.org/10.1016/j.ijcard.2016.02.023>.
- [44] Frommeyer G, Ellermann C, Dechering DG, Kochhäuser S, Bögeholz N, Güner F, et al. Ranolazine and vernakalant prevent ventricular arrhythmias in an experimental whole-heart model of short QT syndrome. *J Cardiovasc Electrophysiol* 2016;27(10):1214–19. Available from: <https://doi.org/10.1111/jce.13029>.
- [45] Sumitomo N, Harada K, Nagashima M, Yasuda T, Nakamura Y, Aragaki Y, et al. Catecholaminergic polymorphic ventricular tachycardia: electrocardiographic characteristics and optimal therapeutic strategies to prevent sudden death. *Heart (Br Card Soc)* 2003;89(1):66–70. Available from: <https://doi.org/10.1136/heart.89.1.66>.
- [46] De Ferrari GM, Dusi V, Spazzolini C, Bos JM, Abrams DJ, Berul CI, et al. Clinical management of catecholaminergic polymorphic ventricular tachycardia the role of left cardiac sympathetic denervation. *Circulation* 2015;131(25):2185–93. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.115.015731>.
- [47] Mohamed U, Gollob MH, Gow RM, Krahn AD. Sudden cardiac death despite an implantable cardioverter-defibrillator in a young female with catecholaminergic ventricular tachycardia. *Heart Rhythm* 2006;3(12):1486–9. Available from: <https://doi.org/10.1016/j.hrthm.2006.08.018>.
- [48] Padfield GJ, Alahmari L, Lieve KVV, Alahmari T, Roston TM, Wilde AA, et al. Flecainide monotherapy is an option for selected patients with catecholaminergic polymorphic ventricular tachycardia intolerant of  $\beta$ -blockade. *Heart Rhythm* 2016;13(2):609–13. Available from: <https://doi.org/10.1016/j.hrthm.2015.09.027>.

- [49] Kaneshiro T, Naruse Y, Nogami A, Tada H, Yoshida K, Sekiguchi Y, et al. Successful catheter ablation of bidirectional ventricular premature contractions triggering ventricular fibrillation in catecholaminergic polymorphic ventricular tachycardia with RyR2 mutation. *Circ: Arrhythmia Electrophysiol* 2012;5(1):14–18. Available from: <https://doi.org/10.1161/CIRCEP.111.966549>.
- [50] Tokioka K, Kusano KF, Morita H, Miura D, Nishii N, Nagase S, et al. Electrocardiographic parameters and fatal arrhythmic events in patients with brugada syndrome: combination of depolarization and repolarization abnormalities. *J Am Coll Cardiol* 2014;63(20):2131–8. Available from: <https://doi.org/10.1016/j.jacc.2014.01.072>.
- [51] Nademanee K, Veerakul G, Chandanamattha P, Chaothawee L, Ariyachaipanich A, Jirasirojanakorn K, et al. Prevention of ventricular fibrillation episodes in brugada syndrome by catheter ablation over the anterior right ventricular outflow tract epicardium. *Circulation* 2011;123(12):1270–9. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.110.972612>.



# Chronic heart failure

*Abbas Zaidi and Parin Shah*

Department of Cardiology, University Hospital of Wales, Cardiff, United Kingdom

## 10.1 Introduction

Heart failure is a clinical syndrome in which patients have symptoms (such as breathlessness, fatigue, and ankle swelling) and signs (tachypnea, pulmonary rales, pleural effusions, raised jugular venous pressure, and peripheral edema) with objective evidence of structural or functional abnormality of the heart at rest [1]. The heart failure syndrome may result from a variety of pathological states when the heart is unable to adequately meet the metabolic demands of the body. Identifying the underlying cause is important in the treatment of heart failure, since precise etiological diagnosis enables specific targeting of therapy.

By far, the most common scenario manifesting with clinical heart failure is an intrinsic abnormality of myocardial function. The most widely utilized index of myocardial function is the left ventricular ejection fraction (LVEF), which is the proportion of the left ventricular end-diastolic volume that is ejected with each beat ( $>55\%$  in normal hearts). Patients with signs and/or symptoms of heart failure and  $\text{LVEF} <40\%$  are classified as having “heart failure with reduced ejection fraction.” Patients who have signs and/or symptoms of heart failure with  $\text{LVEF} >50\%$  are described as having “heart failure with normal ejection fraction,” if there is concomitant evidence of either cardiac dysfunction (e.g., raised natriuretic peptide levels or diastolic dysfunction) or structural heart disease. Recently, the European Society of Cardiology (ESC) has defined a novel subgroup of patients with “heart failure with midrange ejection fraction,” which is defined by an LVEF between  $40\%$  and  $49\%$ , elevated natriuretic peptides, and structural abnormality of the heart [1]. Although there are some inherent limitations to its use as an index of myocardial function, categorization of heart failure patients according to LVEF has proven useful in clinical practice, since these subgroups vary in their etiology and responses to different therapies [2].

## 10.2 Epidemiology

Heart failure affects  $1\%–2\%$  of the adult population in the Western world. The prevalence increases with age, peaking at around  $7\%–8\%$  in those aged more than 75 years [3]. Epidemiological studies suggest that roughly half of the patients with heart failure have reduced LVEF, whereas the other half have normal LVEF [4]. In the United Kingdom the National Heart Failure Audit (2014–15) reported that more than  $70\%$  of the patients hospitalized with acute heart failure had reduced LVEF [5]. Whilst the prevalence of heart failure in men and women is thought to be similar, women are vastly underrepresented in clinical studies, constituting only a quarter of subjects enrolled into the pivotal trials. Female patients with heart failure tend to present at an older age and more often have preserved systolic function. Ethnic variation exists in the incidence of heart failure. In the United States the highest incidence is seen amongst African Americans, with the lowest incidence amongst Chinese Americans. Whilst this is likely to reflect genetic differences to an extent, there are a multitude of confounding variables such as hypertension, diabetes, and socioeconomic factors [6].

### 10.3 Etiology of heart failure

The causes of heart failure can broadly be categorized as (A) diseases of the myocardium, (B) diseases of abnormal cardiac loading, and (C) arrhythmias (Table 10.1). The first category, diseases of the myocardium, includes ischemic heart disease, that is, by far the most common cause of heart failure in Western countries. Primary heart muscle diseases with Mendelian patterns of inheritance also reside within this group. The most important amongst these genetic conditions are hypertrophic cardiomyopathy (typically characterized by left ventricular wall thickening), dilated cardiomyopathy (left ventricular dilatation and impairment), arrhythmogenic cardiomyopathy (predominantly right ventricular dilatation and impairment), restrictive cardiomyopathy (impaired left ventricular relaxation), and left ventricular noncompaction cardiomyopathy (left ventricular hypertrabeculation and impairment). Clinical characteristics, diagnostic criteria, management, and risk stratification of the genetic cardiomyopathies will be the primary focus of this chapter.

**TABLE 10.1** Causes of heart failure.

Causes of heart failure		
1. Diseases of the myocardium		
<b>Ischemic heart disease</b>		
Direct cytotoxicity	Recreational substance abuse	Alcohol, cocaine, amphetamine, anabolic steroids
	Heavy metals	Copper, iron, lead, cobalt
	Medications	Cytotoxic drugs, immunomodulatory drugs, antidepressants, antiarrhythmics, nonsteroidal antiinflammatory drugs, anesthetic agents
	Radiation	Radiotherapy
Immune-mediated and inflammatory damage	Infections	Bacteria, spirochetes, fungi, protozoa, parasites (Chagas disease), rickettsiae, viruses, for example, human immunodeficiency virus, prior myocarditis
	Noninfectious	Lymphocytic/giant cell myocarditis, autoimmune diseases, hypersensitivity and eosinophilic myocarditis
Infiltrative	Malignancy	
	Not malignancy related	Amyloidosis, sarcoidosis, hemochromatosis, glycogen storage disorders, lysosomal storage disorders
Metabolic derangements	Hormonal	Thyroid disease, parathyroid disease, acromegaly, growth hormone deficiency, hypercortisolaemia, Conn's disease, Addison's disease, diabetes, metabolic syndrome, pheochromocytoma, pathologies related to pregnancy
	Nutritional	Deficiency in thiamine, L-carnitine, selenium, iron, phosphates, calcium, complex malnutrition, obesity
Genetic abnormalities		Hypertrophic, dilated, noncompaction, arrhythmogenic, and restrictive cardiomyopathies, muscular dystrophies, mitochondrial diseases, laminopathies
2. Abnormal cardiac loading conditions		
<b>Hypertension</b>		
Valvular and structural cardiac defects	Acquired or congenital	Mitral, aortic, tricuspid and pulmonary valve disease, atrial and ventricular septal defects
	Pericardial	Constrictive pericarditis, pericardial effusion
Pericardial and endomyocardial pathologies	Endomyocardial	Hypereosinophilic syndrome, endomyocardial fibrosis, endocardial fibroelastosis
High-output states		Severe anemia, sepsis, thyrotoxicosis, Paget's disease, arteriovenous fistula, pregnancy
Volume overload		Renal failure, iatrogenic fluid overload
3. Arrhythmias		
Tachyarrhythmias		Atrial and ventricular arrhythmias
Bradyarrhythmias		Sinus node dysfunction, conduction disorders

Adapted from Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). *Eur Heart J* 2016;37:2129–200.

## 10.4 Clinical assessment

Clinical assessment of the patient with signs and/or symptoms of heart failure begins with a comprehensive personal and family history, followed by a thorough physical examination. Clinical investigations are performed for the purposes of diagnosis, risk stratification, and to guide therapy. Routine clinical investigations integrate information from laboratory studies, electrocardiography, multimodality cardiac imaging, and in some cases histopathology. The role of genetic testing will be covered in the following sections of this chapter.

### 10.4.1 Personal history

Typical symptoms such as exertional dyspnea, orthopnea, and paroxysmal nocturnal dyspnea point toward the diagnosis of heart failure. Palpitations may indicate underlying tachyarrhythmias that may have an etiological role (tachycardia-induced cardiomyopathy) or be secondary to myocardial disease. The presence or absence of prior presyncope or syncope should be elicited, since patients with heart failure are at risk of sudden death due to ventricular arrhythmias and conduction system disease. Exertional chest discomfort should prompt investigation for ischemic heart disease. A comprehensive drug history should be taken, including quantification of alcohol consumption and illicit substance misuse, as well as exposure to cardiotoxic drugs such as anthracycline chemotherapeutic agents.

### 10.4.2 Family history

A detailed family history must be elicited in all heart failure patients, particularly if there is a suspicion of a primary genetic cardiomyopathy. Any family history of established cardiomyopathy, or sudden/unexplained deaths at a young age (<40 years old), should be enquired about. A family history of epilepsy, road traffic accidents, or drowning may in some cases indicate an underlying heritable cardiac condition giving rise to ventricular arrhythmias as the mode of death. Postmortem data such as autopsy reports, ideally with both macroscopic and microscopic examination findings, should be sought.

### 10.4.3 Physical examination

Typical physical signs in heart failure are primarily a reflection of fluid congestion and include pulmonary rales, pleural effusions, elevated jugular venous pressure, and peripheral edema. The cardiac apex may be laterally displaced if the left ventricle is dilated. Right ventricular dilatation may be perceived as a parasternal heave. Murmurs may be present if there is underlying valve disease or outflow tract obstruction. Obesity and nicotine stains indicate risk factors for ischemic heart disease. An irregular pulse may indicate underlying atrial fibrillation or other arrhythmias that can promote the development of ventricular dysfunction or have developed as a result of it.

### 10.4.4 Laboratory studies

#### 10.4.4.1 Plasma natriuretic peptides

In recent years, plasma natriuretic peptides, which have a high negative predictive value but low positive predictive value for the diagnosis of heart failure, have been used as a “rule out” test when echocardiography is not immediately available, for example, in the community or general practice setting [7]. The most commonly utilized biomarkers are B-type natriuretic peptide (BNP) and N-terminal pro-BNP (NTproBNP). The normal upper values are 35 and 125 pg/mL for BNP and NTproBNP, respectively, in the nonacute, outpatient setting. In the acute setting the corresponding normal upper values are 100 and 300 pg/mL for BNP and NTproBNP, respectively [1]. Natriuretic peptides are also useful for monitoring response to treatments, and for early detection of incipient acute decompensation [8].

#### 10.4.4.2 Cardiomyopathy screen

A variety of laboratory investigations should be considered in patients with unexplained heart failure or cardiomyopathy, according to the particular clinical background and phenotype (Table 10.2).

**TABLE 10.2** Laboratory investigations in heart failure.

Investigation	Associated conditions
Renal and liver function, complete blood count, bone profile, thyroid stimulating hormone	Various
Creatine kinase	Mitochondrial disease, storage disorders, myotonic dystrophy
Serum iron, ferritin	Hemochromatosis, secondary iron overload, for example, transfusion-dependent patients
Serum free light chains, urine and plasma protein immunofixation	Amyloidosis (restrictive cardiomyopathy)
Urinary protein	Fabry disease, amyloidosis
$\alpha$ -Galactosidase A	Fabry disease
Plasma lactic acid	Mitochondrial diseases
Urine myoglobin	Mitochondrial diseases
Autoantibody screen	Various
Viral titers (coxsackievirus, echovirus, influenza, human immunodeficiency virus, <i>Borrelia burgdorferi</i> , Chagas serology)	Dilated cardiomyopathy
Plasma thiamine	Dilated cardiomyopathy
Urinary catecholamines	Hypertrophic and dilated cardiomyopathy
Serum angiotensin converting enzyme	Sarcoidosis (causing phenocopy of dilated or arrhythmogenic cardiomyopathy)

## 10.4.5 Electrocardiography

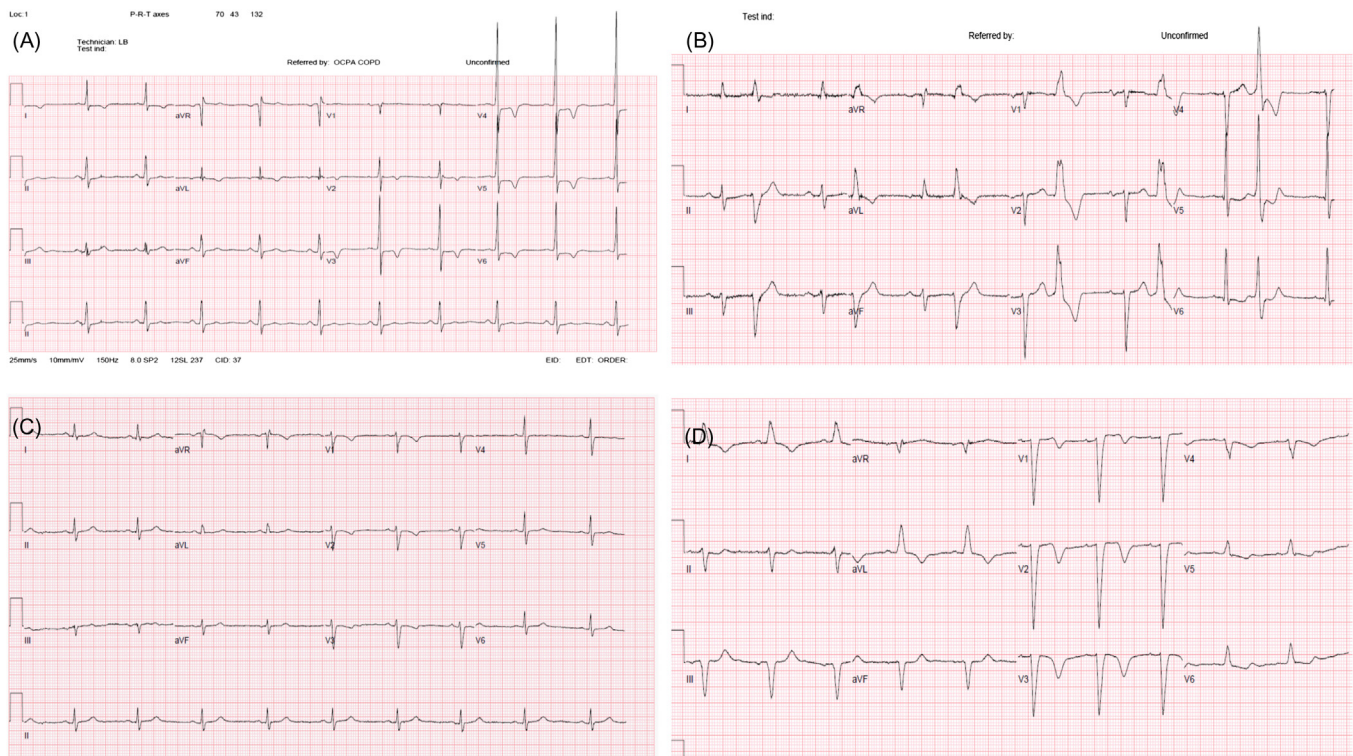
### 10.4.5.1 Resting electrocardiography

The 12-lead electrocardiogram (ECG) is a pivotal investigation in the diagnosis and risk stratification of heart failure patients. Nonspecific markers of cardiac structural abnormality or myocardial pathology include Q-waves, axis deviation, ST-segment deviation, and repolarization changes such as T-wave inversion. Prolongation of the QRS interval, particularly with left bundle branch block morphology and exceeding 150 ms, is a marker of cardiac mechanical dyssynchrony that may benefit from cardiac resynchronization therapy (CRT) [9]. Reduced QRS complex amplitudes may indicate myocardial infiltration with amyloid (cardiac amyloidosis) or fat (arrhythmogenic cardiomyopathy). The genetic cardiomyopathies are characterized by several other more specific markers. Whilst left ventricular hypertrophy by voltage criteria and associated repolarization changes (strain pattern) can be seen with any pathology causing left ventricular wall thickening or cavity dilatation, deep T-wave inversion in leads V3/4–V6 is typical of the apical form of hypertrophic cardiomyopathy. T-wave inversion in leads V1–V3, or small positive deflections at the end of the QRS complex (epsilon waves) in the same leads, is suggestive of arrhythmogenic cardiomyopathy [10]. Ventricular preexcitation is indicated by a short PR interval and slurred upstroke to the QRS complex (delta wave). This finding in a patient with a cardiomyopathy phenotype should raise suspicion for an underlying inheritable metabolic disorder such as Fabry disease. Typical electrocardiographic findings in patients with different causes of heart failure are demonstrated in Fig. 10.1.

### 10.4.5.2 Ambulatory electrocardiography

Arrhythmias are common in patients with heart failure. Ambulatory monitoring should be used in patients reporting palpitations, syncope, or presyncope. Asymptomatic arrhythmias are also common and can carry significant morbidity and mortality. For example, atrial fibrillation is highly prevalent in heart failure patients, and anticoagulation is usually indicated due to high thromboembolic risk. The presence of complex ventricular arrhythmias should prompt consideration of prophylactic defibrillator implantation. In patients with diagnostic uncertainty, such as an asymptomatic family member of a patient with hypertrophic cardiomyopathy, or an athlete with left ventricular hypertrophy of uncertain etiology, the presence of ventricular arrhythmias is a strong indicator of underlying disease [11]. Bradycardias may also be seen in heart failure patients, as a side effect of





**FIGURE 10.1** 12-Lead electrocardiograms from patients with different causes of heart failure. (A) Hypertrophic cardiomyopathy. There is left ventricular hypertrophy by voltage criteria, with high QRS complex amplitudes in the lateral precordial leads V4–V6. There is abnormal repolarization in leads V2–V6, seen as inverted T-waves and ST-segment depression. (B) Dilated cardiomyopathy. ST-segment depression is present in the lateral leads V5–V6. Frequent ventricular extrasystoles are seen, which may be a harbinger of sustained ventricular arrhythmias predisposing to sudden death. (C) Arrhythmogenic cardiomyopathy with predominant right ventricular involvement. QRS voltages are globally reduced due to fibrofatty infiltration of the myocardium. There is T-wave inversion in leads V1–V3, which overlie the right ventricle. (D) Ischemic cardiomyopathy. Leads V1–V3 reveal Q-waves, loss of R-wave, persistent ST-segment elevation, and T-wave inversion, in keeping with extensive, transmural anterior myocardial infarction. There is also left bundle branch block; hence this patient may benefit from cardiac resynchronization therapy by implantation of a biventricular pacemaker.

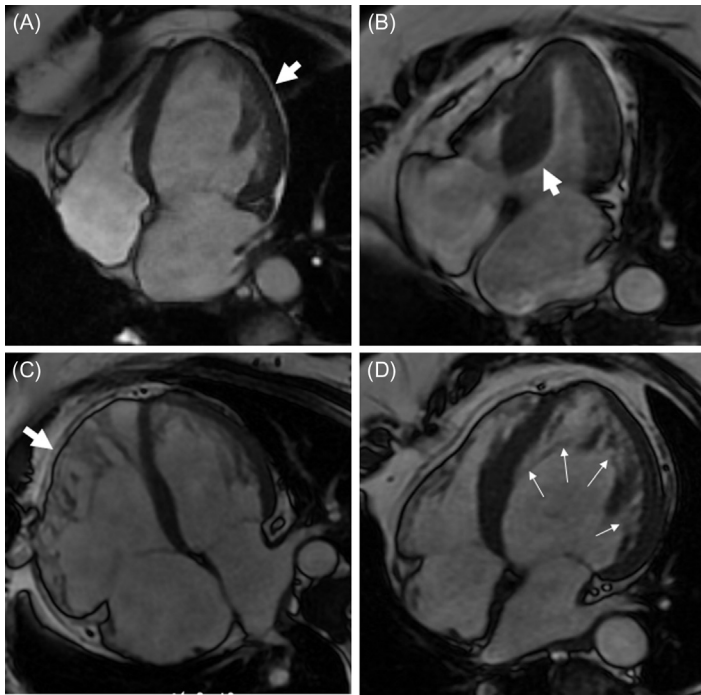
therapies such as beta-blockers, or due to infiltration of the conduction system in diseases such as sarcoidosis and amyloidosis. Premature conduction disorders such as complete heart block in a young patient with a dilated cardiomyopathy phenotype should prompt consideration of specific genotypes such as lamin A/C mutations [12].

### 10.4.6 Imaging studies

All patients with suspected heart failure should undergo a plain chest radiograph and transthoracic echocardiogram. Cardiac magnetic resonance (CMR) imaging is increasingly utilized in the diagnosis of heart failure, due to its unique strengths in tissue characterization. Other imaging modalities may include computerized tomography (CT) and nuclear studies such as positron emission tomography (PET). Fig. 10.2 shows characteristic morphological findings in the common genetic cardiomyopathies as visualized by CMR.

#### 10.4.6.1 Plain chest radiography

The plain chest X-ray is a quick, inexpensive, and widely available tool, which still has significant diagnostic ability in patients with suspected heart failure. Cardiac enlargement may be seen in the form of an increased cardiothoracic ratio. Signs of decompensation include pleural effusions and pulmonary congestion (perihilar alveolar shadowing, upper lobe vascular diversion, Kerley B-lines, and fluid in the horizontal fissure). Pointers toward specific etiologies may occasionally be seen, for example, pericardial calcification in constrictive pericarditis, or bilateral hilar lymphadenopathy in sarcoidosis.



**FIGURE 10.2** Genetic cardiomyopathies. All images shown are cardiac magnetic resonance apical four chamber views. (A) Dilated cardiomyopathy. The left ventricle (arrow) is severely dilated and impaired. (B) Hypertrophic cardiomyopathy, with asymmetric interventricular septal hypertrophy (arrow). (C) Arrhythmogenic cardiomyopathy with predominant right ventricular (arrow) involvement. (D) Left ventricular noncompaction cardiomyopathy, characterized by prominent left ventricular hypertrabeculation (arrows).

#### 10.4.6.2 Echocardiography

Transthoracic echocardiography forms the cornerstone of diagnosis, risk stratification, and therapeutic monitoring in all types of heart failure. It is almost universally available, is noninvasive, and provides exquisite functional information due to extremely high temporal resolution. The LVEF derived by echocardiography is the most widely studied index of cardiac performance in heart failure, and much of the guidance regarding therapy relies on this parameter. Its major limitation is that LVEF represents primarily radial systolic function. It is now increasingly recognized that long axis cardiac dysfunction is a more sensitive indicator of myocardial pathology, and that LVEF may only deteriorate at an advanced stage in the pathological process [13]. Nonetheless, key decision-making criteria, including thresholds for device therapies such as implantable defibrillators, still rely on LVEF. In heart failure patients with midrange (LVEF 40%–49%) or preserved ejection fraction (LVEF > 50%), echocardiography is critical in demonstrating impaired cardiac relaxation, primarily by means of Doppler assessment of left ventricular filling during diastole.

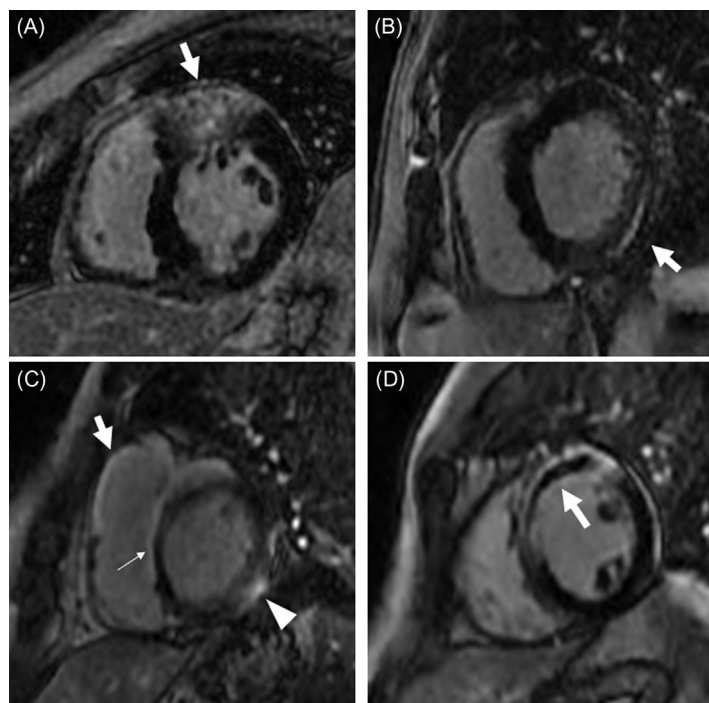
Echocardiography also provides vital morphological information regarding the etiology of heart failure. Patients with ischemic heart disease will typically exhibit segmental myocardial dysfunction, with or without wall thinning, in the distribution of the diseased coronary vessel/s. Echocardiography remains the optimal imaging modality for demonstrating the mechanism, severity, and potential therapeutic options for patients with heart failure due to valve disease. Hypertensive heart disease, a common cause of heart failure with preserved ejection fraction, will usually reveal concentric left ventricular hypertrophy with reduced cavity size and impaired diastolic filling. Hypertrophic cardiomyopathy most commonly manifests with predominant hypertrophy of the interventricular septum (asymmetric septal hypertrophy pattern), although variants with concentric or predominantly apical hypertrophy are seen [14]. The mitral valve and subvalvar apparatus are often abnormal in hypertrophic cardiomyopathy, with apical displacement of the papillary muscles, and elongation of the mitral valve leaflets. Around one-third of patients with hypertrophic cardiomyopathy exhibit left ventricular outflow tract obstruction at rest, due to a combination of septal hypertrophy and the aforementioned mitral valve abnormalities. A hypertrophic cardiomyopathy phenotype with concentric hypertrophy and associated right ventricular hypertrophy should raise suspicion for a metabolic disorder such as Fabry or Danon disease, particularly when presenting in infancy, or if there is evidence of ventricular preexcitation on the ECG [15]. Restrictive cardiomyopathy is characterized by severely impaired diastolic cardiac filling with or without left ventricular hypertrophy, which relies heavily on Doppler echocardiography for diagnosis. Restrictive cardiomyopathy presenting at a young age should also raise suspicion for a metabolic disorder. Dilated cardiomyopathy is characterized by global left ventricular (and often right ventricular) dilatation, wall thinning, and systolic dysfunction. However, there are no specific features on echocardiography, which help to differentiate primary



dilated cardiomyopathy from that caused by secondary factors such as excessive alcohol consumption [16]. Prominent left ventricular trabeculation in a dilated and dysfunctional heart may indicate left ventricular noncompaction cardiomyopathy. This is considered to be a disorder in which the inner of the two layers of the myocardium fails to “compact” normally during development [17]. Arrhythmogenic cardiomyopathy, in its most common form, results in dilatation of the right ventricle accompanied by regional wall motion abnormalities or aneurysms. Current diagnostic criteria for this condition focus on quantification of right ventricular dimensions and functional parameters, including the right ventricular outflow tract by echocardiography [10]. Current diagnostic criteria are conspicuous for their absence of reference to the left ventricle, which is now recognized to be involved to some extent in a significant proportion of cases. This has led to a change in terminology such that arrhythmogenic cardiomyopathy with predominant right, predominant left, or biventricular involvement is preferred over the previous nomenclature (arrhythmogenic right ventricular cardiomyopathy).

#### 10.4.6.3 Cardiac magnetic resonance imaging

The development of CMR imaging has heralded a new era in our understanding of the myocardial changes underlying different causes of heart failure. This is a consequence of the unique tissue characterization capabilities of MRI, which exploits the fact that different types of tissue demonstrate characteristic relaxation kinetics after excitation with a radiofrequency pulse within a strong magnetic field. These differences can be amplified by the administration of gadolinium-based contrast. The technique of “late gadolinium enhancement” imaging has allowed for the first time a reliable method for the visualization, localization, and quantification of myocardial fibrosis. Different patterns of fibrosis are recognized in different etiologies of heart failure [18]. Subendocardial fibrosis is usually indicative of ischemic myocardial damage. The extent of subendocardial fibrosis is strongly associated with the likelihood of functional recovery after coronary revascularization (viability). Subepicardial fibrosis is a typical finding if the previous myocarditis is the cause of heart failure. Band-like, mid-wall fibrosis is typically seen in nonischemic dilated cardiomyopathy. Hypertrophic cardiomyopathy is often accompanied by patchy fibrosis in the region/s of maximal hypertrophy, often the interventricular septum. The presence of fibrosis can help to differentiate hypertrophic cardiomyopathy from phenocopies such as hypertensive heart disease (usually little or no fibrosis) and Fabry disease (inferolateral wall fibrosis). Arrhythmogenic cardiomyopathy may reveal right ventricular fibrosis, and left ventricular fibrosis is also increasingly recognized. Infiltrative conditions such as sarcoidosis and amyloidosis are characterized by patchy scar that does not conform to any particular coronary territory. Fig. 10.3 shows late gadolinium enhancement patterns in patients with different causes of heart failure.



**FIGURE 10.3** Tissue characterization of different causes of heart failure by late gadolinium enhancement cardiac magnetic resonance imaging. All images shown are short axis slices through the left and right ventricles. Healthy myocardium appears black, whereas regions of myocardial fibrosis appear white. (A) Hypertrophic cardiomyopathy, with asymmetric interventricular septal hypertrophy and patchy fibrosis in the region of maximal hypertrophy (arrow). (B) Dilated cardiomyopathy, with linear mid-wall fibrosis in the inferolateral segment (arrow). (C) Arrhythmogenic cardiomyopathy with biventricular involvement. There is widespread fibrosis in the right ventricular free wall (broad arrow), interventricular septum (thin arrow), and inferolateral segment of the left ventricle (arrowhead). (D) Ischemic cardiomyopathy. There is transmural fibrosis confined to the distribution of the left anterior descending coronary artery, indicating a full thickness anterior myocardial infarction. The black region within the white region of fibrosis indicates microvascular obstruction. Full thickness fibrosis and microvascular obstruction in the setting of acute infarction indicates a low probability of functional recovery in that myocardial segment after revascularization (nonviability).

In addition to its tissue characterization qualities, CMR allows excellent visualization of cardiac structure and function even in cases where echocardiographic windows are suboptimal (such as obese patients or those with chronic lung disease) and is the gold standard for quantification of biventricular volumes and ejection fraction. The right ventricle is particularly well seen by CMR in comparison with echocardiography, and CMR should ideally be performed in any patient with suspected arrhythmogenic cardiomyopathy. CMR also allows visualization of myocardial inflammation and edema in conditions such as myocarditis. Finally, several relatively novel parametric mapping techniques are rapidly transitioning into everyday clinical practice and have demonstrated extremely high sensitivity for phenotypically mild or preclinical myocardial disease [19]. These techniques include T1 mapping (diffuse fibrosis), T2 mapping (myocardial edema), and T2\* mapping (myocardial iron overload).

#### **10.4.6.4 Other imaging modalities**

Nuclear imaging techniques, such as PET, have proven particularly useful in demonstrating the presence of inflammatory myocardial disease such as sarcoidosis. Very recent PET studies are also challenging our understanding of metabolic heart diseases such as Fabry cardiomyopathy, in which widespread myocardial inflammation can be seen by PET, much earlier than the development of overt cardiac hypertrophy or focal fibrosis on CMR [20]. CT is particularly useful for defining precise anatomy, since its strength lies in its superior spatial resolution. CT is useful for delineating coronary anatomy and excluding significant atheroma as an alternative to invasive angiography (e.g., in younger heart failure patients with a low likelihood of coronary disease). CT is also the investigation of choice for demonstrating pericardial calcification in suspected constrictive pericarditis. In practice, the majority of heart failure patients will undergo invasive coronary angiography that remains the gold standard for the diagnosis of ischemic heart disease. A proportion of these will undergo left and right heart catheterization with invasive measurement of cardiac hemodynamics, particularly if there is pulmonary hypertension, or if there is suspicion of intracardiac shunting or pericardial constriction.

#### **10.4.7 Exercise testing**

It is increasingly acknowledged that resting assessment of cardiac function provides an incomplete reflection of a patient's clinical condition, since patients with cardiac disease usually experience symptoms during exertion [21]. Hence in patients with exertional dyspnea due to suspected cardiac failure, exercise testing, ideally with concomitant echocardiography, should be performed if resting studies are unrevealing. Specific features to assess during exertion include worsening of valvular regurgitation, increasing pulmonary artery pressures, development of cardiac dyssynchrony, or impairment of diastolic filling. Patients with hypertrophic cardiomyopathy may develop "dynamic" outflow tract or intracavitary obstruction only during exercise. Provocation of cardiac arrhythmias or abnormal blood pressure responses during exercise testing should also be noted. Cardiopulmonary exercise testing is useful in defining whether exercise limitation is due to cardiac disease or other pathology in patients with multiple comorbidities such as concomitant chronic lung disease.

#### **10.4.8 Cardiac biopsy**

Histopathological examination of the cardiac tissue remains the gold standard investigation for the diagnosis of many primary and secondary myocardial diseases. This may be postmortem at autopsy, or in living subjects by means of catheter-based endomyocardial biopsy [22]. The diagnostic criteria for myocarditis have traditionally relied heavily on biopsy to demonstrate inflammatory infiltrates and leukocytes, with or without fibrosis, and for isolation and amplification of viral RNA. Myocardial infiltration may be demonstrated in conditions such as amyloidosis (apple-green birefringence with Congo red stain), eosinophilic myocarditis, and sarcoidosis (noncaseating granulomas). In cases of heart failure with left ventricular hypertrophy where the etiology is unclear, the demonstration of myocyte disarray is pathognomonic for sarcomeric hypertrophic cardiomyopathy rather than phenocopies such as hypertensive heart disease. There may in addition be fibrosis and small vessel disease in hypertrophic cardiomyopathy. Arrhythmogenic cardiomyopathy is characterized by fibrofatty myocardial replacement at biopsy. This can help to differentiate from other conditions in which there may be right ventricular enlargement and fatty infiltration without fibrosis, such as obesity.

## 10.5 Genetic testing

### 10.5.1 Limitations of clinical assessment

Whilst the approach to the patient presenting with heart failure must involve a comprehensive clinical assessment, there are limitations to this strategy. The signs and symptoms of heart failure are in general not disease specific and are a poor guide to etiology. Similarly, whilst there are some relatively specific electrocardiographic signs such as the epsilon wave in arrhythmogenic cardiomyopathy, the majority of ECG anomalies in heart failure syndromes are nonspecific markers of myocardial disease (such as Q-waves or voltage criteria for ventricular hypertrophy) and do not point toward a particular cause. Advances in imaging techniques have improved our understanding of the myocardial pathology underlying different conditions causing heart failure. Nonetheless, there is still a significant phenotypic overlap between conditions with vastly different etiology, heritability, and prognosis. For example, physiological left ventricular hypertrophy in response to regular exercise may closely resemble phenotypically mild hypertrophic cardiomyopathy, the latter carrying the risk of sudden death during exertion. Even the traditional diagnostic gold standard of cardiac biopsy is not without its shortcomings, since it is an invasive procedure that carries a risk of morbidity and has the potential for false negative results if there is patchy myocardial disease. The clinical approach also has limitations when considering the surveillance of relatives of individuals affected by an inheritable cardiomyopathy, given the highly variable penetrance of these conditions. Traditionally, familial evaluation has involved longitudinal, potentially lifelong, surveillance of relatives by means of symptom review, electrocardiography, and imaging tests. The financial burden of maintaining entire families under long-term clinical follow-up is significant, in addition to the uncertainty and anxiety resulting for these individuals. The use of genetic testing in conjunction with clinical assessment offers a means by which to address some of these limitations.

### 10.5.2 The genetics of acquired heart failure

There is a growing body of literature indicating a genetic influence on the development, progression, and response to treatment of acquired forms of heart failure such as ischemic cardiomyopathy. For example, genome-wide association studies have implicated the ubiquitin-specific protease gene (USP3) with the development of heart failure in patients of European ancestry, and the leucine-rich, immunoglobulin-like domain (LRIG3) in individuals of African ancestry [23]. Data from candidate gene studies have revealed differential responses to heart failure therapy in patients with different beta adrenoceptor and angiotensin converting enzyme polymorphisms [24,25]. Whilst this is an important area for ongoing research, it is not currently ready for clinical application, and the remainder of this section will focus on the inheritable primary cardiomyopathies.

### 10.5.3 The genetic basis of the inheritable cardiomyopathies

The following section will deal with the approach to the patient with hypertrophic, dilated, arrhythmogenic, restrictive, and noncompaction cardiomyopathies, which usually demonstrate a Mendelian pattern of inheritance. The mode of inheritance is most commonly autosomal dominant. Notable exceptions to this include Fabry disease cardiomyopathy (X-linked) and mitochondrial phenocopies of hypertrophic and dilated cardiomyopathies. There is usually a high degree of genetic heterogeneity in the primary cardiomyopathies [26]. For example, more than 400 mutations have been associated with the development of hypertrophic cardiomyopathy. Penetrance in the genetic cardiomyopathies is generally age related, and relatively few genotype-phenotype correlations have been characterized to date. Table 10.3 shows the genetic characteristics and common mutations associated with the most frequent inheritable causes of heart failure.

By far, the most common mutations associated with hypertrophic cardiomyopathy are those encoding the myosin heavy chain (MYH7) and myosin-binding protein C (MYBPC3). Each of these accounts for approximately 40% of genotype-positive cases, followed by troponin T (TNNT2) and troponin I (TNNI3) mutations (5% each) [27]. Specific phenotypes in association with hypertrophic cardiomyopathy should prompt testing for particular genes, for example, ventricular preexcitation (PRKAG2), or evidence of systemic metabolic disease (LAMP2 mutations in Danon disease, or GLA mutations leading to alpha galactosidase deficiency in Fabry disease) [28]. A large number of genes are associated with the development of dilated cardiomyopathy, with lamin A/C (LMNA) mutations traditionally considered to be the most prevalent (around 6% of familial cases) [29]. Recent interest, however, has focused on the role of titin truncating mutations, which have been reported in up to 25% of familial cases of dilated cardiomyopathy [30]. The most common mutations associated with

**TABLE 10.3** Key genes and encoded proteins associated with genetic cardiomyopathies.

Condition	Gene	Protein
Hypertrophic cardiomyopathy	MYH7	Myosin heavy chain
	MYBPC3	Myosin-binding protein C
	TNNT2	Cardiac troponin T
	TNNI3	Cardiac troponin I
	TPM1	Tropomyosin 1 alpha chain
Dilated cardiomyopathy	LMNA	Lamin A/C
	MYH7	Myosin heavy chain
	MYPN	Myopalladin
	TNNT2	Cardiac troponin T
	SCN5A	Sodium channel
	RBM20	RNA-binding protein 20
	TMPO	Thymopoetin
	TTN	Titin
	DSG2	Desmoglein 2
Arrhythmogenic cardiomyopathy	PKP2	Plakophilin 2
	DSP	Desmoplakin
	DSC2	Desmocollin 2
	JUP	Junctional plakoglobin
	TMEM43	Transmembrane protein 43
	RYR2	Ryanodine receptor 2
	MYH7	Myosin heavy chain
Left ventricular noncompaction cardiomyopathy	MYBPC3	Myosin-binding protein C
	TNNT2	Cardiac troponin T
	TNNI3	Cardiac troponin I
	TPM1	Tropomyosin 1 alpha chain
	LMNA	Lamin A/C

arrhythmogenic cardiomyopathy involve the desmosomal proteins plakophilin 2, desmoglein 2, desmoplakin, junctional plakoglobin, and desmocollin 2. Several nondesmosomal genes, including those encoding transforming growth factor beta 3 (TGFB3) and the ryanodine receptor, are also implicated [31]. Many of the genes associated with hypertrophic and dilated cardiomyopathies, including MYH7 and TNNT2, are also implicated in the pathogenesis of left ventricular noncompaction and restrictive cardiomyopathies [32,33]. Genetic overlap between different inheritable cardiomyopathies is common and also exists between cardiomyopathies and ion channel diseases. For example, mutations in the SCN5A gene are observed in the cardiac ion channelopathy, Brugada syndrome, as well as in arrhythmogenic cardiomyopathy [34]. Understanding how mutations in one gene can lead to vastly different phenotypes remains a substantial challenge in the field of cardiac genetics.

#### 10.5.4 Indications for genetic testing in cardiomyopathies

##### 10.5.4.1 Diagnostic confirmation and prognostication in clinically suspected cases

In cases with a definite clinical cardiomyopathy phenotype, additional genetic testing is generally not required for diagnostic confirmation alone. In certain specific circumstances where there is diagnostic uncertainty after

comprehensive clinical evaluation, for example, in an athlete with mild left ventricular hypertrophy of uncertain etiology, genetic testing may be of additional benefit. Demonstration of a mutation associated with hypertrophic cardiomyopathy could be lifesaving in such a case, since the athlete could be advised not to engage in competitive sport. Genetic testing may also be warranted in cases where the cardiomyopathy phenotype is suspicious for a particular mutation that requires specific precautions or therapies. For example, a patient with a hypertrophic cardiomyopathy phenotype and lateral wall late gadolinium enhancement on CMR would raise suspicion for Fabry disease. Identification of a GLA mutation in such a case could facilitate prompt commencement of enzyme replacement therapy, which may in turn retard disease progression. Genetic testing for a particularly high-risk genotype may be warranted in some cases on prognostic grounds. For example, positive testing for lamin A/C mutations in dilated cardiomyopathy, or TNN2 mutations in hypertrophic cardiomyopathy, might strengthen the argument for early prophylactic defibrillator implantation in patients with “red flags” such as premature conduction system disease or a strong family history of sudden death.

#### **10.5.4.2 Predictive testing of family members**

In cases where a disease-causing mutation has been identified in an index case, predictive testing for that specific mutation may be offered to asymptomatic family members who have not yet developed the phenotype [26]. This may include relatives of index cases whose genotype has been identified postmortem using molecular autopsy [35]. More controversially, it may also permit prenatal genetic testing of unborn family members. Predictive genetic testing for cardiomyopathies has the benefit of removing uncertainty, allows unaffected family members to be discharged from longitudinal clinical surveillance, and permits heightened awareness for symptomatic and phenotypic changes in genotype-positive individuals.

### **10.5.5 Genetic testing techniques in heart failure**

Advances in clinical genetics have made testing more widely available, sensitive, and cost-effective. Until recently, genetic testing for cardiomyopathies has relied on targeted Sanger sequencing of a small number of genes. A typical Sanger sequencing-based hypertrophic cardiomyopathy panel would have included testing for approximately 10 of the most common disease-causing mutations. The development of next-generation sequencing has allowed a much larger number of genes to be interrogated at the same time, which may in turn facilitate greater diagnostic yield [36]. A modern hypertrophic cardiomyopathy panel, for example, typically includes around 40 genes, including selected noncoding disease-causing variants. Targeted sequencing has the advantage of being highly specific and relatively inexpensive. Next-generation sequencing has also permitted increased use of whole-exome sequencing, which attempts to capture all protein-encoding regions in the genome, and whole-genome sequencing that also assesses intronic regions which may modify disease expression [37]. The use of whole-exome and whole-genome sequencing is currently limited by high costs. In addition, the coverage of the exome may not be as complete as with targeted panels. In practice, the mainstay of genetic testing in cardiomyopathies remains the targeted gene panel. In families where a disease-causing mutation has been identified, predictive testing of other family members generally only requires targeted testing for that specific gene.

### **10.5.6 The cardiac genetics multidisciplinary team**

Patients with inheritable cardiac conditions predisposing to heart failure should be assessed within the setting of a heart muscle disease clinic, by physicians with training and expertise in inherited cardiac conditions. In cases where referral for genetic testing is being considered, the case should be reviewed by a cardiac genetics multidisciplinary team. This should include cardiologists with expertise in inherited cardiac conditions and advanced imaging, a pediatric cardiologist, a clinical geneticist, genetics nurses, and genetic counselors [38]. The input of the genetic counselor can be vital in terms of researching the family pedigree, educating patients about the condition, discussing the risks and benefits of genetic testing, and providing psychological support when needed. The team should meet on a regular basis to review and update cases, and robust local databases should be developed and maintained.

### **10.5.7 Limitations of genetic testing in heart failure**

There are certain limitations in the genetic testing of heart failure patients, which the clinician and patient should be aware of before testing is offered. Whilst the financial cost of testing has come down over the years,



and will continue to do, it is still relatively expensive. The process of referral, counseling, and testing is often lengthy, and there may be considerable anxiety and uncertainty for patients during this period. The increased availability of genetic testing, and techniques such as whole-exome sequencing, has compounded the problem posed by variants of uncertain significance. It may be several years before the pathogenicity of any particular variant can be confidently confirmed or discounted. A positive genetic result may have implications for life insurance, mortgages, and sports participation, even in asymptomatic, phenotypically normal family members of index cases. As such, predictive testing of children in particular remains highly controversial. Patients should understand that disease expression in cardiomyopathies is highly variable, a positive test is often a poor guide to the rate and extent of gene expression and clinical outcome, and that there is in most cases no therapeutic intervention that can alter phenotypic expression. There may be significant psychological consequences of a positive result for the individual concerned and their wider family. Conversely, it is important that patients and families understand that genetic testing may be negative even when the phenotype is positive, due to our incomplete understanding of the genetic causation of these conditions. In such cases, longitudinal clinical surveillance of family members is still required. Finally, the phenomenon of “survivor guilt” may occur in genotype negative individuals within a genetically positive cardiomyopathy family [39].

## 10.6 Management

The goal of treatment in heart failure patients is to improve symptoms, avoid hospital admissions, and to improve prognosis (including prevention of sudden cardiac death). Treatments that improve functional capacity are also important to consider, since heart failure is a chronic disease which significantly affects quality of life. There are various treatments, both pharmacological and nonpharmacological, which have been demonstrated to improve symptoms and/or prognosis in patients with heart failure.

### 10.6.1 Pharmacological therapies

During the past three decades, several major clinical trials of neurohormonal antagonists (angiotensin converting enzyme inhibitors [40], angiotensin receptor blockers [41], mineralocorticoid receptor antagonists [42], and beta-blockers [43]) have shown benefit in improving symptoms, reducing hospitalization, and improving prognosis in heart failure patients. When used individually their benefits are modest, but in combination these drugs may reduce 2-year mortality by 50%–60% [44]. Very recently, a new therapeutic class, the angiotensin receptor neprilysin inhibitor (ARNI), has been introduced into clinical practice. This drug class acts on the renin–angiotensin–aldosterone system and the neutral endopeptidase system. The ARNI sacubitril–valsartan has proven superior to enalapril in preventing heart failure admissions, cardiovascular mortality, and all-cause mortality [45]. Sacubitril–valsartan treatment is therefore recommended in heart failure patients who remain symptomatic despite optimal medical therapy [1].

In contrast, none of the pharmacotherapies that are beneficial in heart failure patients with reduced LVEF have shown mortality or morbidity benefit in those with normal LVEF [46]. Considering the prevalence, outcomes, future projections, and lack of effective therapies, heart failure with normal LVEF represents one of the largest unmet needs in cardiovascular medicine.

Specific pharmacological therapies may be beneficial in the genetic cardiomyopathies. The principle objective is symptom reduction and improvement of functional capacity, since there is little evidence that pharmacotherapy reduces sudden death risk in these conditions. Beta-blockers are effective in ameliorating symptomatic left ventricular outflow tract obstruction in hypertrophic cardiomyopathy [14]. Verapamil or diltiazem may be used if beta-blockers are contraindicated, and the sodium channel blocker disopyramide may be added for persistent symptoms. Drug-refractory patients may benefit from surgical ventricular septal myectomy. An alternative is cardiac catheter-based septal ablation, in which pure ethanol is injection into a septal perforator artery in order to induce a localized iatrogenic septal myocardial infarction, thereby reducing the degree of obstruction. Beta-blockers and amiodarone are effective in suppressing ventricular arrhythmias in arrhythmogenic cardiomyopathy, although again there appears to be no associated mortality benefit [47]. Catheter ablation of ventricular tachycardia may be used in arrhythmogenic cardiomyopathy patients with frequent symptomatic arrhythmias, although recurrence rates are high due to the patchy and progressive nature of myocardial involvement. Atrial fibrillation is common in all of the genetic cardiomyopathies, due to elevated cardiac filling pressures leading to



atrial dilatation. Thromboembolic risk in hypertrophic cardiomyopathy is high, and anticoagulation is warranted for atrial fibrillation even in the absence of other thromboembolic risk factors. Left ventricular noncompaction has been associated with an increased risk of thromboembolism in some studies, and anticoagulation may be considered even in the absence of atrial fibrillation.

### 10.6.2 Nonsurgical device therapies and risk stratification for sudden death

In heart failure patients with LVEF <35% and a broad QRS complex (>130 ms), CRT should be considered. CRT is a form of permanent pacing involving simultaneous stimulation of the left and right ventricles, in an attempt to augment cardiac performance and improve hemodynamics. CRT is effective in improving symptoms and has been shown to reduce both morbidity and mortality [48]. CRT should be considered in heart failure patients who are symptomatic despite optimal medical treatment.

The most serious complication of heart failure is sudden cardiac death. Secondary prevention therapy with an implantable cardioverter–defibrillator (ICD) is generally indicated in heart failure patients surviving an episode of hemodynamically unstable, sustained ventricular arrhythmia. Much effort has been directed toward the identification of high-risk individuals in whom lifestyle modifications, pharmacotherapy, and prophylactic (primary prevention) ICD implantation might prevent sudden death. Optimal treatment with guideline-recommended therapies (beta-blockers, mineralocorticoid receptor antagonists, sacubitril/valsartan, and CRT) reduces this risk to an extent. Antiarrhythmic drugs such as amiodarone are effective in suppressing tachyarrhythmias but ineffective in reducing overall mortality in heart failure [49]. The ESC recommends prophylactic ICD implantation in patients with symptomatic heart failure (ischemic and nonischemic etiologies) and severely reduced LVEF ( $\leq 35\%$ ) despite optimal medical therapy [1,50]. In practice, since the indications for CRT and ICD therapy overlap significantly, a combined “CRT-D” device (CRT—defibrillator) is implanted in many cases. There are currently no indications for prophylactic ICD therapy in heart failure patients with preserved ejection fraction.

Several of the inheritable cardiomyopathies, including hypertrophic and arrhythmogenic cardiomyopathies, have separate criteria for prophylactic ICD use. For hypertrophic cardiomyopathy, current ESC guidance advocates estimation of the patient’s 5-year sudden death risk according to their age, maximal left ventricular wall thickness, left atrial size, degree of left ventricular outflow obstruction, prior history of syncope, ventricular arrhythmias, and family history of sudden death. A total risk score  $\geq 6\%$  over 5 years is considered an indication for ICD therapy [14,51]. North American guidance differs slightly, with blunted blood pressure response during exercise testing also considered a high-risk feature [52]. There is also evidence that certain hypertrophic cardiomyopathy mutations may be associated with a higher risk of sudden death, for example, troponin T mutations that may present with sudden death even in the absence of significant LV hypertrophy [53].

Major risk factors for sudden death in arrhythmogenic cardiomyopathy include prior ventricular arrhythmias, severe right ventricular or left ventricular impairment, and syncope. Other less strongly validated risk factors include male gender, proband status, provokable arrhythmias at electrophysiological study, and extensive T-wave inversion [54]. As with hypertrophic cardiomyopathy, specific genotypes may be relevant to sudden death risk. Carriers of multiple mutations in the same desmosomal gene, or mutations in two or more genes, appear to be at greater arrhythmic risk than those with a single mutation. Multivariable analysis has shown the presence of multiple desmosomal gene mutations to be an independent predictor of lifetime major arrhythmic events with a hazard ratio of 2.3 [55]. In addition, the rare nondesmosomal TMEM43-p. S358L missense mutation has been shown to be fully penetrant and highly lethal amongst a founder population in Newfoundland [56].

### 10.6.3 Exercise

In the past, it was commonly assumed that rest was beneficial in heart failure patients. However, clinical trials of exercise training in heart failure patients have demonstrated unequivocal benefits in terms of symptom relief, with a suggestion that training might even improve prognosis. All current guidelines now firmly recommend regular physical activity and structured exercise training for patients with heart failure [57].

In certain circumstances, exercise restriction may be warranted. Patients with genetic cardiomyopathies are generally at increased risk of sudden death during intense exercise [58]. This is due to interactions between the abnormal underlying substrate (fibrotic myocardium) and exercise-induced electrolyte and acid–base disturbances, adrenaline surges, and microvascular ischemia, which may in turn lead to the development of lethal ventricular arrhythmias. Until very recently, all competitive sports participation had been prohibited for individuals

with established cardiomyopathy. The latest ESC guidance has taken a more lenient approach, allowing participation for athletes with phenotypically mild, asymptomatic genetic cardiomyopathies without high-risk features such as unexplained syncope or arrhythmias. The exception to this is arrhythmogenic cardiomyopathy, in which all competitive sports are prohibited due to the high arrhythmic risk, as well as the established role of exercise in promoting phenotypic development. Similarly, the new guidelines permit athletes who are genotype positive but phenotypically negative for cardiomyopathy to participate in all competitive sports, again with the exception of arrhythmogenic cardiomyopathy [59].

## References

- [1] Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JG, Coats AJ, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). *Eur Heart J* 2016;37:2129–200.
- [2] Butler J, Fonarow GC, Zile MR, Lam CS, Roessig L, Schelbert EB, et al. Developing therapies for heart failure with preserved ejection fraction: current state and future directions. *JACC Heart Fail* 2014;2:97–112.
- [3] Redfield MM, Jacobsen SJ, Burnett JC Jr, et al. Burden of systolic and diastolic ventricular dysfunction in the community: appreciating the scope of the heart failure epidemic. *J Am Med Assoc* 2003; 289:194–202.
- [4] Hogg K, Swedberg K, McMurray J. Heart failure with preserved left ventricular systolic function; epidemiology, clinical characteristics, and prognosis. *J Am Coll Cardiol* 2004;43:317–27.
- [5] National Heart Failure Audit. Annual report 2014/15, <<https://www.ucl.ac.uk/nicor/audits/heartfailure/reports>>; 2016 [accessed 25.04.17].
- [6] Bahrami H, Kronmal R, Bluemke DA, Olson J, Shea S, Liu K, et al. Differences in the incidence of congestive heart failure by ethnicity. The multi-ethnic study of atherosclerosis. *Arch Intern Med* 2008;168:2138–45.
- [7] Maisel A, Mueller C, Adams K, Anker SD, Aspromonte N, Cleland JGF, et al. State of the art: using natriuretic peptide levels in clinical practice. *Eur J Heart Fail* 2008;10:824–39.
- [8] Stienen S, Salah K, Moons AH, Bakx AL, van Pol P, Kortz RAM, et al. NT-proBNP (N-terminal pro-B-type natriuretic peptide)-guided therapy in acute decompensated heart failure: PRIMA II randomized controlled trial (can NT-ProBNP-guided therapy during hospital admission for acute decompensated heart failure reduce mortality and readmissions?). *Circulation* 2018;137:1671–83.
- [9] Cleland JG, Daubert JC, Erdmann E, Freemantle N, Gras D, Kappenberger L, et al. Cardiac resynchronization-heart failure (CARE-HF) study investigators. The effect of cardiac resynchronization on morbidity and mortality in heart failure. *N Engl J Med* 2005;352:1539–49.
- [10] Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, et al. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the task force criteria. *Circulation* 2010;121:1533–41.
- [11] Sharma S, Merghani A, Mont L. Exercise and the heart: the good, the bad, and the ugly. *Eur Heart J* 2015;36:1445–53.
- [12] Fatkin D, MacRae C, Sasaki T, Wolff MR, Porcu M, Frenneaux M, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* 1999;341:1715–24.
- [13] Thavendiranathan P, Poulin F, Lim KD, Plana JC, Woo A, Marwick TH. Use of myocardial strain imaging by echocardiography for the early detection of cardiotoxicity in patients during and after cancer chemotherapy: a systematic review. *J Am Coll Cardiol* 2014;63:2751–68.
- [14] Elliott PM, Anastakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). *Eur Heart J* 2014;35:2733–79.
- [15] Yousef Z, Elliott PM, Cecchi F, Escoubet B, Linhart A, Monserrat L, et al. Left ventricular hypertrophy in Fabry disease: a practical approach to diagnosis. *Eur Heart J* 2013;34:802–8.
- [16] Weintraub RG, Semsarian C, Macdonald P. Dilated cardiomyopathy. *Lancet* 2017;390:400–14.
- [17] Towbin JA, Lorts A, Jefferies JL. Left ventricular non-compaction cardiomyopathy. *Lancet* 2015;386:813–25.
- [18] Captur G, Manisty C, Moon JC. Cardiac MRI evaluation of myocardial disease. *Heart* 2016;102:1429–35.
- [19] Swoboda PP, McDiarmid AK, Page SP, Greenwood JP, Plein S. Role of T1 mapping in inherited cardiomyopathies. *Eur Cardiol* 2016;11:96–101.
- [20] Spinelli L, Imbriaco M, Nappi C, Nicolai E, Giugliano G, Ponsiglione A, et al. Early cardiac involvement affects left ventricular longitudinal function in females carrying  $\alpha$ -galactosidase A mutation: role of hybrid positron emission tomography and magnetic resonance imaging and speckle-tracking echocardiography. *Circ Cardiovasc Imaging* 2018;11:e007019.
- [21] Corrà U, Agostoni PG, Anker SD, Coats AJS, Crespo Leiro MG, de Boer RA, et al. Role of cardiopulmonary exercise testing in clinical stratification in heart failure. A position paper from the Committee on Exercise Physiology and Training of the Heart Failure Association of the European Society of Cardiology. *Eur J Heart Fail* 2018;20:3–15.
- [22] Cooper LT, Baughman KL, Feldman AM, Frustaci A, Jessup M, Kuhl U, et al. The role of endomyocardial biopsy in the management of cardiovascular disease: a scientific statement from the American Heart Association, the American College of Cardiology, and the European Society of Cardiology Endorsed by the Heart Failure Society of America and the Heart Failure Association of the European Society of Cardiology. *Eur Heart J* 2007;28:3076–93.
- [23] Smith NL, Felix JF, Morrison AC, Demissie S, Glazer NL, Loehr LR, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet* 2010;3:56–266.
- [24] Terra SG, Hamilton KK, Pauly DF, Lee CR, Patterson JH, Adams KF, et al. Beta1-adrenergic receptor polymorphisms and left ventricular remodeling changes in response to beta-blocker therapy. *Pharmacogenet Genomics* 2005;15:227–34.

- [25] McNamara DM, Holubkov R, Postava L, Janosko K, MacGowan GA, Mathier M, et al. Pharmacogenetic interactions between angiotensin converting enzyme inhibitor therapy and the angiotensin-converting enzyme deletion polymorphism in patients with congestive heart failure. *J Am Coll Cardiol* 2004;44:2019–26.
- [26] Charron P, Arad M, Arbustini E, Basso C, Bilinska Z, Elliott P, European Society of Cardiology Working Group on Myocardial and Pericardial Diseases, et al. Genetic counselling and testing in cardiomyopathies: a position statement of the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. *Eur Heart J* 2010;31:2715–26.
- [27] Maron BJ, Maron MS, Semsarian C. Genetics of hypertrophic cardiomyopathy after 20 years: clinical perspectives. *J Am Coll Cardiol* 2012;60:705–15.
- [28] Wicks EC, Elliott PM. Genetics and metabolic cardiomyopathies. *Herz* 2012;37:598–610.
- [29] Pankuweit S. Lamin A/C mutations in patients with dilated cardiomyopathy. *Eur Heart J* 2018;39:861–3.
- [30] Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012;366:619–28.
- [31] Gandjbakhch E, Redheuil A, Pousset F, Charron P, Frank R. Clinical diagnosis, imaging, and genetics of arrhythmogenic right ventricular cardiomyopathy/dysplasia: JACC state-of-the-art review. *J Am Coll Cardiol* 2018;72:784–804.
- [32] Arbustini E, Favalli V, Narula N, Serio A, Grasso M. Left ventricular noncompaction: a distinct genetic cardiomyopathy? *J Am Coll Cardiol* 2016;68:949–66.
- [33] Muchtar E, Blauwet LA, Gertz MA. Restrictive cardiomyopathy: genetics, pathogenesis, clinical manifestations, diagnosis, and therapy. *Circ Res* 2017;121:819–37.
- [34] Corrado D, Zorzi A, Cerrone M, Rigato I, Mongillo M, Bauce B, et al. Relationship between arrhythmogenic right ventricular cardiomyopathy and Brugada syndrome: new insights from molecular biology and clinical implications. *Circ Arrhythm Electrophysiol* 2016;9.
- [35] Cann F, Corbett M, O'Sullivan D, Tennant S, Hailey H, Grieve JH, et al. Phenotype-driven molecular autopsy for sudden cardiac death. *Clin Genet* 2017;91:22–9.
- [36] Parikh VN, Ashley EA. Next-generation sequencing in cardiovascular disease: present clinical applications and the horizon of precision medicine. *Circulation* 2017;135:406–9.
- [37] Seidelmann SB, Smith E, Subrahmanyam L, Dykas D, Abou Ziki MD, Azari B, et al. Application of whole exome sequencing in the clinical diagnosis and management of inherited cardiovascular diseases in adults. *Circ Cardiovasc Genet* 2017;10.
- [38] Arscott P, Caleshu C, Kotzer K, Kreykes S, Kruisselbrink T, Orland K, et al. A case for inclusion of genetic counselors in cardiac care. *Cardiol Rev* 2016;24:49–55.
- [39] Care M, Chauhan V, Spears D. Genetic testing in inherited heart diseases: practical considerations for clinicians. *Curr Cardiol Rep* 2017;19:88.
- [40] The CONSENSUS Trial Study Group. Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). *N Engl J Med* 1987;316:1429–35.
- [41] Cohn JN, Tognoni G. A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. *N Engl J Med* 2001;345:1667–75.
- [42] Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. *N Engl J Med* 1999;341:709–17.
- [43] Packer M. Effect of carvedilol on the morbidity of patients with severe chronic heart failure: results of the Carvedilol Prospective Randomized Cumulative Survival (COPERNICUS) study. *Circulation* 2002;106:2194–9.
- [44] Cleland JG, Clark AL. Delivering the cumulative benefits of triple therapy to improve outcomes in heart failure: too many cooks will spoil the broth. *J Am Coll Cardiol* 2003;42:1234–7.
- [45] McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, et al. Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N Engl J Med* 2014;371:993–1004.
- [46] Cleland JG, Pellicori P, Dierckx R. Clinical trials in patients with heart failure and preserved left ventricular ejection fraction. *Heart Fail Clin* 2014;10:511–23.
- [47] Mazzanti A, Ng K, Faragli A, Maragna R, Chiodaroli E, Orphanou N, et al. Arrhythmogenic right ventricular cardiomyopathy: clinical course and predictors of arrhythmic risk. *J Am Coll Cardiol* 2016;68:2540–50.
- [48] Bristow MR, Saxon LA, Boehmer J, Krueger S, Kass DA, De Marco T, et al. Cardiac-resynchronization therapy with or without an implantable defibrillator in advanced chronic heart failure. *N Engl J Med* 2004;350:2140–50.
- [49] Bardy GH, Lee KL, Mark DB, Poole JE, Packer DL, Boineau R, et al. Amiodarone or an implantable cardioverter–defibrillator for congestive heart failure. *N Engl J Med* 2005;352:225–37.
- [50] Moss AJ, Hall WJ, Cannom DS, Daubert JP, Higgins SL, Klein H, et al. Improved survival with an implanted defibrillator in patients with coronary disease at high risk for ventricular arrhythmia. Multicenter Automatic Defibrillator Implantation Trial Investigators. *N Engl J Med* 1996;335:1933–40.
- [51] O'Mahony C, Jichi F, Pavlou M, Monserrat L, Anastasakis A, Rapezzi C, et al. A novel clinical risk prediction model for sudden cardiac death in hypertrophic cardiomyopathy (HCM risk-SCD). *Eur Heart J* 2014;35:2010–20.
- [52] Gersh BJ, Maron BJ, Bonow RO, Dearani JA, Fifer MA, Link MS, et al. 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation* 2011;124:2761–96.
- [53] Pasquale F, Syrris P, Kaski JP, Mogensen J, McKenna WJ, Elliott P. Long-term outcomes in hypertrophic cardiomyopathy caused by mutations in the cardiac troponin T gene. *Circ Cardiovasc Genet* 2012;5:10–17.
- [54] Calkins H, Corrado D, Marcus F. Risk stratification in arrhythmogenic right ventricular cardiomyopathy. *Circulation* 2017;136:2068–82.
- [55] Bhonsale A, Groeneweg JA, James CA, Dooijes D, Tichnell C, Jongbloed JD, et al. Impact of genotype on clinical course in arrhythmogenic right ventricular dysplasia/cardiomyopathy-associated mutation carriers. *Eur Heart J* 2015;36:847–55.
- [56] Merner ND, Hodgkinson KA, Haywood AF, Connors S, French VM, Drenckhahn JD, et al. Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene. *Am J Hum Genet* 2008;82:809–21.

- [57] Piepoli MF, Conraads V, Corra U, Dickstein K, Francis DP, Jaarsma T, et al. Exercise training in heart failure: from theory to practice. A consensus document of the Heart Failure Association and the European Association for Cardiovascular Prevention and Rehabilitation. *Eur J Heart Fail* 2011;13:347–57.
- [58] Corrado D, Basso C, Rizzoli G, Schiavon M, Thiene G. Does sports activity enhance the risk of sudden death in adolescents and young adults? *J Am Coll Cardiol* 2003;42:1959–63.
- [59] Pelliccia A, Solberg EE, Papadakis M, Adami PE, Biffi A, Caselli S, et al. Recommendations for participation in competitive and leisure time sport in athletes with cardiomyopathies, myocarditis, and pericarditis: position statement of the Sport Cardiology Section of the European Association of Preventive Cardiology (EAPC). *Eur Heart J* 2019;40:19–33.

# Molecular pathophysiology of systemic hypertension

*Tatiana Garofalidou<sup>1</sup> and Patricia B. Munroe<sup>1,2</sup>*

<sup>1</sup>William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom <sup>2</sup>National Institutes for Health Research, Barts Cardiovascular Biomedical Research Centre, Queen Mary University of London, London, United Kingdom

## 11.1 Introduction

Hypertension is a highly prevalent condition in the adult population. It affects one billion people worldwide, representing 31% of the entire adult population [1]. More specifically, it is estimated to occur in one in three young adults, which increases to 60% of the population over 60 years and three out of four people aged over 70 [2]. It is a major cause of morbidity and mortality due to its association with cardiovascular disease, stroke, renal impairment, and peripheral vascular disease [3]. Worldwide, 51% of mortality is due to cerebrovascular disease and 45% due to ischemic heart disease attributable to high blood pressure (BP) [4]. The Global Burden of Disease study has indicated that hypertension remains the largest single risk factor for global all-cause mortality, contributing to approximately 10 million deaths and 212 million lost healthy life years per year [5]. Therefore hypertension is a major global public health challenge, and successful prevention and treatment are essential in promoting longevity and reducing disease burden worldwide.

Essential hypertension, also known as primary hypertension, is defined as high BP, without a clear single identifiable cause. Hypertension itself is defined by most guidelines as a systolic BP (SBP) > 140 mmHg, and/or a diastolic BP (DBP) > 90 mmHg. For every 20 mmHg increase of SBP and 10 mmHg of DBP the increased risk for stroke and heart disease doubles and, thus, from a clinical perspective, hypertension can be better defined as the BP level at which BP-lowering treatment results in a significant clinical benefit.

Systemic hypertension is caused by several interrelated factors, which contribute to the development of elevated BP in the systemic arteries. These include not only nonmodifiable risk factors, such as age, sex, and ethnic background, but also multiple modifiable risk factors. The most common modifiable risk factors are excessive dietary sodium and poor dietary potassium intake, obesity, physical inactivity, smoking, increased alcohol intake, and physiological stress [6,7]. However, hypertension is a complex disease, which means its pathogenesis is ultimately determined by an interplay between genetic alterations and environmental/lifestyle factors [5].

## 11.2 Blood pressure regulation—key systems

### 11.2.1 Endothelium

BP is defined as the product of cardiac output, which is the volume of blood pumped per minute, and systemic vascular resistance, determined by both intravascular volume as well as neuronal and humoral factors. These include the renin–angiotensin aldosterone system (RAAS), the sympathetic nervous system, the immune



system, and natriuretic peptides acting on the endothelium. Sodium ( $\text{Na}^+$ ) plays an important role in BP levels, as a high serum sodium concentration leads to water retention and, thus, an increase in BP. In normotensive individuals with a high serum sodium concentration, compensatory mechanisms are in place to maintain constant BP levels. These include a decrease in the peripheral and renal vascular resistance, as well as an increase in nitric oxide (NO) production, which is a potent vasodilator. However, in the absence of a normal NO response, BP increases and, therefore, endothelial dysfunction is a predictor for the development of salt-sensitive hypertension. The endothelium is also responsible for releasing a variety of other vasodilators including prostacyclin and endothelial-derived hyperpolarizing factor, as well as vasoconstrictors such as endothelin-1 and locally acting angiotensin II. The endothelium therefore plays an important role in the pathogenesis of hypertension.

### 11.2.2 Renin–angiotensin aldosterone system

The RAAS system is a crucial regulator BP homeostasis, particularly through its mechanism of action in the kidney. Renin is synthesized and released by the juxtaglomerular apparatus cells in the kidney in response to a low sodium concentration, increased vasodilation, an increased renal sympathetic activity, and a decreased renal afferent perfusion pressure. It is responsible for cleaving angiotensinogen to angiotensin I, which is in turn converted to angiotensin II by the angiotensin-converting enzyme. Angiotensin II stimulates angiotensin II receptors type 1 and leads to smooth muscle cell contraction, systemic vasoconstriction and hence an increased vascular resistance, and an increased sodium reabsorption in the proximal convoluted tubule (PCT) through the stimulation of the  $\text{Na}^+/\text{H}^+$  exchanger 3, the  $\text{Na}^+/\text{HCO}_3^-$  cotransporter 1, and the  $\text{Na}^+/\text{K}^+$ -ATPase. It also increases aldosterone synthesis and release from the adrenal glomerulosa. Activation of the angiotensin II receptors type 2 has the opposite effects, leading to vasodilation and natriuresis.

### 11.2.3 Natriuretic peptides

Natriuretic peptides are also important in sodium and BP homeostasis and interact with the RAAS, through inhibiting renin and aldosterone release, thus leading to an increase in BP levels. Atrial natriuretic peptide and brain natriuretic peptide are released by the stretched atria and ventricles and lead to vasodilation, natriuresis, and by extension a decreased fluid volume. They also inhibit the reabsorption of sodium by the kidneys through inhibiting the  $\text{Na}^+/\text{K}^+$ -ATPase and the sodium–glucose cotransporter in the PCT, and the epithelial sodium channel (ENaC) in the distal tubule.

### 11.2.4 Sympathetic nervous system

Hypertension leads to vessel-wall distension. Baroreceptors are mechanoreceptors, present in the carotid sinus, which respond to arterial stretch due to high BP, and stimulate the vasomotor center of the brain to reduce the efferent sympathetic activity and decrease BP. It has been found that in hypertensive individuals, there is an increased release of, as well as increased sensitivity to, norepinephrine, accompanied by an increased responsiveness to stress and a decreased baroreceptor sensitivity.

### 11.2.5 Immune system

Inflammation and the immune system have also been found to play a role in BP regulation and the pathogenesis of hypertension. Inflammation is associated with the release of many mediators, including cytokines, reactive oxygen species, and NO. Cytokines are responsible for the formation of a neointima and the enhancement of vascular fibrosis, leading to increased peripheral vascular resistance, while they also increase the synthesis of angiotensinogen and angiotensin II, promoting sodium and water reabsorption in the kidney. Both innate and adaptive immunity play a role in the development of hypertension.

---

## 11.3 Genetics of hypertension

The majority of patients (90%–95%) present with essential hypertension (EH), as a result of complex gene-environment interactions. The familial aggregation of BP combined with the frequent occurrence of a positive family history in hypertensive individuals suggests an underlying genetic predisposition to hypertension. The



genetic basis of hypertension was first established through family and twin studies, which enabled the assessment of the heritability of BP. Studies have estimated the heritability of BP to be between 48% and 60% for SBP and between 34% and 67% for DBP [8]. However, family and twin heritability studies do not permit the identification of the specific genetic variations that may play a role in the development of hypertension or, by extension, the underlying mechanism(s) by which genetic differences influence BP.

Over the past decade, genomic tools have rapidly advanced, allowing the genotyping of millions of single nucleotide polymorphisms (SNPs) in an efficient, reliable, and cost-effective manner. Genome-wide association studies (GWAS) utilize dense sets of SNPs, which are evenly spaced across the entire human genome, and genotyping is usually done on cases (individuals with a disease) and controls (people without disease). Association testing is performed to determine any differences in allele frequencies between the groups, any associated SNPs provide regions of the genome-containing genes which influence disease risk [3]. The GWAS approach has enabled the discovery of hundreds of loci associated with BP and hypertension and is providing new insights into previously undiscovered biochemical or molecular pathways underlying BP control. The findings from these studies permit new disease classifications based on molecular mechanisms rather than the phenotype alone, and these loci represent as potential drug targets for both prevention and management of hypertension. The BP GWAS studies have identified and validated 901 BP loci to date, together explaining 27% of the estimated heritability of BP, while functional analyses of the GWAS findings have shed light on new biological pathways for BP regulation, thus holding some potential of improving disease prevention [9].

All forms of hypertension have some underlying genetic basis. Nevertheless, there is substantial variation in the genetic contribution across the different forms of hypertension [3]. In the general population, BP measurements follow a normal distribution, where hypertensive individuals have inherited a collection of genetic variants that, in combination with environmental risk factors, predispose them being in the higher end of the distribution. The majority of the SNPs associated with BP are common in the population and have relatively small effect sizes on the BP phenotype, typically up to 1 mmHg.

There are several rare Mendelian/monogenic forms of hypertension; here a single gene mutation is responsible for the pathogenesis of the condition. Eleven monogenic forms of hypertension have been described, which present with distinctive cophenotypes. These include Gordon's syndrome (pseudohypoaldosteronism type 2) (PHA2), Liddle's syndrome, familial hyperaldosteronism; type 1—glucocorticoid remediable aldosteronism (GRA), and types 2, 3, and 4, congenital adrenal hyperplasia (CAH) with 11 $\beta$ -hydroxylase deficiency (type 4), CAH with 17 $\alpha$ -hydroxylase deficiency (type 5), apparent mineralocorticoid excess, hypertension associated with brachydactyly, and Geller syndrome, also known as hypertension exacerbated by pregnancy.

Next generation sequencing technologies have permitted the identification of the causative genes for monogenic forms of hypertension, and gene discovery is the key for the targeted management of affected individuals. The main characteristics of all 10 monogenic forms of hypertension are shown in Table 11.1. The clinical, genetic, and pathophysiological aspects of Gordon's syndrome, Liddle's syndrome, CAH (types 4 and 5), and GRA will be discussed in detail in this chapter.

## 11.4 Monogenic forms of systemic hypertension

### 11.4.1 Gordon's syndrome

#### 11.4.1.1 Case report: a 52-year-old man with hypertension and hyperkalemia was presented

Tom, a 52-year-old man, with hypertension and nonspecific symptoms including nausea, fatigue, and muscle weakness was presented. Laboratory results reveal hyperkalemia, and a hyperchloremic metabolic acidosis, despite normal glomerular filtration rate. His brother and son are identically affected. Plasma renin activity (PRA) is low, while plasma aldosterone is normal. Jack is prescribed a mineralocorticoid for his hypertension and hyperkalemia, without successful reduction of BP. Further treatment with a thiazide diuretic corrects his BP and the serum electrolytes. Jack is diagnosed with Gordon's syndrome on the basis of his hypertension, hyperkalemia, and hyperchloremic metabolic acidosis, despite normal kidney function as well as his positive family history of the symptoms.

#### 11.4.1.2 Definition

Gordon's syndrome, also known as PHA2 (OMIM:145260) is a rare autosomal dominant disorder, characterized by salt-dependent hypertension and hyperkalemia despite a normal glomerular filtration rate as a result of impaired potassium excretion. The age of onset is variable, with some affected individuals being diagnosed

**TABLE 11.1** Monogenic forms of hypertension.

Condition	Phenotype	Inheritance	Chromosome	Gene	Frequency in general population
Gordon's syndrome; pseudohypoaldosteronism type 2	Hypertension; ↑K <sup>+</sup> ; metabolic acidosis	AD	12p13	<i>WNK1</i>	Rare
		AD	17q21	<i>WNK4</i>	Rare
		AD	2q36	<i>CUL3</i>	Rare
		AD and AR	5q31	<i>KLHL3</i>	Rare
Liddle's syndrome	Hypertension; ↓K <sup>+</sup> ; metabolic alkalosis; ↓PRA; ↓aldosterone	AD	12p13.31	<i>SCNN1A</i>	Rare
		AD	16p12.2	<i>SCNN1B</i>	Rare
		AD	16p12.2	<i>SCNN1G</i>	Rare
Congenital adrenal hyperplasia type 4—11β-hydroxylase deficiency	Hypertension; other variable features; virilization; ↓K <sup>+</sup>	AR	8q	<i>CYP11B1</i>	1/100,000 births, 5%–8% cases
Congenital adrenal hyperplasia type 5—17α-hydroxylase deficiency	Hypertension; delayed sexual development	AR	10q24.32	<i>CYP17A1</i>	Very rare
Familial hyperaldosteronism type 1 (Glucocorticoid remediable aldosteronism)	Hypertension; other variable features	AD	8q24.3	<i>CYP11B1/CYP11B2 chimera</i>	Rare
Familial hyperaldosteronism type 3	Hypertension, ↓K <sup>+</sup> ; ↑aldosterone, other features	AD	11q	<i>KCNJ5</i>	Very rare
Familial hyperaldosteronism Type 4				<i>CACNA1H</i>	Rare
Apparent mineralocorticoid excess	Hypertension; ↓K <sup>+</sup> , ↓↓aldosterone, metabolic alkalosis, ↓renin	AR	16q	<i>HSD11B2</i>	Very rare
Hypertension and brachydactyly	Severe salt independent hypertension; brachydactyly	AD	12p	<i>PDE3A</i>	Very rare
Geller syndrome; Hypertension exacerbated by pregnancy	Early hypertension; exacerbated by pregnancy	AD	4q31.23	<i>NR3C2</i>	One pedigree

AD, Autosomal dominant; AR, autosomal recessive; K, potassium.

in infancy or childhood and others in adulthood. Hyperkalemia usually occurs first with hypertension developing later in life. Metabolic acidosis also occurs due to impaired urinary H<sup>+</sup> excretion, while hyperchloremia, hypercalciuria, and suppressed PRA are variable-associated findings. Individuals with these metabolic abnormalities may also present with nonspecific symptoms such as nausea, vomiting, fatigue, and muscle weakness. To date, more than 180 individuals and families affected by Gordon's syndrome have been reported [10].

#### 11.4.1.3 Genetics

Gordon's syndrome presents with genetic heterogeneity and PHA2 loci have been mapped in different families to five different chromosomes. PHA2A (OMIM: 145260) to chromosome 1q31-q42, PHA2B (OMIM: 614491) to 17q21.2, PHA2C (OMIM: 614492) to 12p13.33, PHA2D (OMIM: 614495) to 5q31.2 and PHA2E (OMIM: 614496) to 2q36.2. Therefore mutations in several genes can be responsible for causing Gordon's syndrome. These include mutations in the *WNK1* and *WNK4* genes, leading to mutations in two members of the serine/threonine kinases and causing PHA2C and PHA2B, respectively. Mutations in the *KLHL3* and the *CUL3* genes are responsible for causing PHA2D and PHA2E, respectively, while the genetic cause of PHA2A remains unknown. It has been shown that PHA2 caused by mutations in the *WNK1* gene presents a less severe phenotype compared to PHA2 caused by mutations in the *WNK4* and *KLHL3* genes. All three types are less severely affected compared to PHA2 (PHA2E) caused by dominant mutations in the *CUL3* gene [11].

#### 11.4.1.3.1 WNK genes

Wilson et al. identified two mutations in the *WNK1* gene responsible for PHA2C in a 10-member kindred segregating PHA2. A 41-kb deletion in intron 1 of *WNK1* gene, as well as a 22-kb deletion in the same intron, previously identified by Disse-Nicodeme et al., were found to increase *WNK1* expression and were not present in the unaffected family members [12]. Using bioinformatic tools, a paralog of *WNK1* was identified as the *WNK4* gene, located on chromosome 17 in a region that had been previously associated with PHA2B. Further examination of the *WNK4* gene in PHA2B kindreds discovered four missense mutations that cosegregated with the disease. Three of these mutations (Glu562Lys, Asp564Ala, and Gln565Glu) lie in exon 7, in an amino acid sequence which is highly conserved among all members of the human WNK kinases. The fourth mutation identified (Arg1185Cys) lies within exon 17 at a residue conserved among *WNK4*, 1 and 2. None of these mutations were found in the unaffected family members [12].

#### 11.4.1.3.2 KLHL3 gene

Boyden et al. studied 52 PHAII kindreds and identified both dominant and recessive mutations in the *KLHL3* gene, which resulted in PHA2D. The recessive mutations (Trp470Ter, Phe322Cys, Arg240Ter, Arg336Ile, and Tyr557Cys) were located along the encoded protein, while the dominant mutations (Ser410Leu, Arg528His, and Ser433Asn) showed marked clustering. They suggested that the dominant mutations led to abnormal binding of the *KLHL3* protein to *CUL3* or other substrates [11]. Louis-Dit-Picard et al. identified missense mutations in the *KLHL3* gene among affected individuals from 16 out of 45 families with hypertension associated with hyperkalemia. The mutations were present both in heterozygosity (Ala398Val, Asn529Lys, Pro426Leu) or in homozygosity, with the recessive cases being diagnosed earlier and the dominant ones having a more severe phenotype [13].

#### 11.4.1.3.3 CUL3 gene

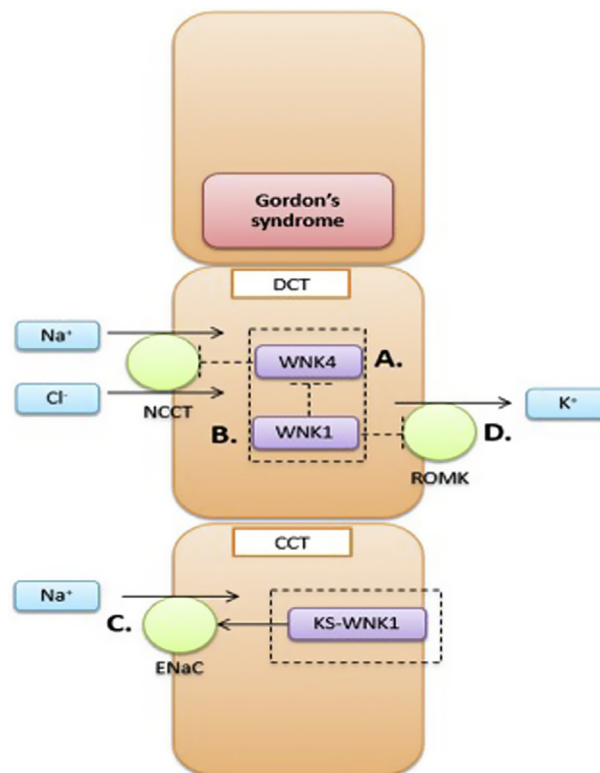
Boyden et al. also identified that mutations in the *CUL3* gene also segregated in PHA2 cases, including both de novo and autosomal dominant forms (PHA2E). Of the 52 kindreds studied, 17 had mutations in the *CUL3* gene, 8 of which were de novo, thus indicating that these mutations are pathogenic. All 17 mutations were located in sites involved in exon 9 splicing [11].

### 11.4.1.4 Pathophysiology

#### 11.4.1.4.1 WNK kinases

The WNK kinases are a four-member family of serine/threonine kinases. Protein kinases are enzymes which regulate the activity of other proteins through phosphorylating specific amino acid residues (serine, threonine, or tyrosine), thus inducing a conformational change from an inactive to an active form of target proteins [14]. They play an important role in signal transduction pathways, to regulate the structure and function of the target proteins in a variety of cellular functions, including cell proliferation, metabolism, apoptosis, gene expression, and many more [15]. The catalytic domain of most protein kinases is highly conserved. The characteristic feature of WNK kinases is that the catalytic lysine residue is located in subdomain I instead of subdomain II, like most other kinases—hence the name “with-no-lysine kinases” (WNK; [16]).

WNK kinases are present in many tissues and cell lines and are therefore pleiotropic, carrying out a variety of functions such including cell growth, neurotransmission, and solute transport. In the kidney the WNKs regulate three major sodium transporters including the sodium potassium chloride cotransporter type 2, the sodium chloride cotransporter (NCCT), and the ENaC, thus playing an important role in renal salt reabsorption as well as potassium and hydrogen ion excretion. More specifically, they are thought to act as “molecular switches” that coordinate the activity of these channels, based on physiological conditions [17]. Normally, *WNK4* is responsible for inhibiting NCCT, preventing sodium and chloride reabsorption in the DCT. Mutations in *WNK4*, therefore, lead to unrestrained sodium and chloride reabsorption leading to the hypertensive phenotype of PHA2. Unlike *WNK4*, *WNK1* does not directly affect sodium reabsorption. It suppresses the inhibitory effect of *WNK4*, thus increasing activity of NCCT. In PHA2, overexpression of the full-length “long” isoform of the WNK transcripts in the kidney, L-*WNK1*, leads to further inhibition of the *WNK4* protein and, thus, enhanced sodium reabsorption, explaining the pathogenesis of hypertension [18]. The second *WNK1* transcript expressed in the kidney is the “short” kidney-specific isoform KS-*WNK1*. It was demonstrated that KS-*WNK1* was induced by aldosterone in a kidney cell line, and overexpression of it led to the enhanced sodium transport via ENaC. Therefore as hyperkalemia in patients with PHA2 leads to aldosterone release, *WNK1* could also contribute directly to sodium reabsorption in the cortical collecting duct (CCD) via the overexpression of the KS-*WNK1* transcript, acting



**FIGURE 11.1** Effects of mutations in WNK kinase genes in the kidney and the hypertensive phenotype in Gordon's syndrome.

- A. Normally, WNK4 inhibits the NCCT channel in the distal convoluted tubule preventing  $\text{Na}^+$  and  $\text{Cl}^-$  reabsorption. Mutations in *WNK4* lead to unrestrained  $\text{Na}^+$  and  $\text{Cl}^-$  reabsorption, leading to the hypertensive phenotype of PHA2.
- B. WNK1 inhibits the inhibitory action of WNK4, enhancing the activity of NCCT. In PHA2, overexpression of the long isoform of the WNK1 transcripts in the kidney, L-WNK1, leads to further inhibition of WNK4, and, therefore, increased  $\text{Na}^+$  reabsorption.
- C. Overexpression of the short WNK1 isoform, KS-WNK1, in the kidney leads to enhanced  $\text{Na}^+$  transport via ENaC in the common convoluted tubule. KS-WNK1 expression is induced by aldosterone, and as hyperkalemia in PHA2 patients leads to aldosterone release, KS-WNK1 could also contribute to the hypertensive phenotype.
- D. The hyperkalemia observed in PHA2 patients could be due to reduced  $\text{K}^+$  excretion by ROMK due to inhibition of ROMK by L-WNK1.

*DCT*, Distal convoluted tubule; *ENaC*, epithelial sodium channel; *NCCT*, thiazide-sensitive  $\text{Na}^+-\text{Cl}^-$  cotransporter; *ROMK*, rectifying outer medullary potassium channel.

directly on ENaC [19]. Vidal-Petiot et al. in 2013 showed in a mouse model that a *WNK1* mutation increased the expression of the renal outer medullary potassium (K) channel (ROMK), which is responsible for potassium excretion in the distal nephron. Therefore the hyperkalemia observed in PHA2 patients could be due to reduced potassium excretion as a result of ROMK inhibition by L-WNK1 [20]. An overview of mutations in WNK kinase genes causing PHAI and PHAII is illustrated in Fig. 11.1.

#### 11.4.1.4.2 KLHL3 and CUL3 proteins

The KLHL3 protein, encoded by the *KLHL3* gene, is part of the ubiquitin-proteasome system. It is a major component of the E3 ubiquitin ligase, a complex which is responsible for the ubiquitination of damaged or excess proteins, so they can be identified and degraded by proteasomes. More specifically, its role involves identifying the target proteins and attaching the E3 ubiquitin ligase complex to them. Target proteins include the WNK1 and WNK4 proteins; therefore, mutations in the *KLHL3* gene impair the ubiquitination and downregulation of WNK1 and WNK4 kinases. This leads to the activation of a phosphorylation cascade which increases the activity of sodium channels and sodium chloride channels at the DCT, including ENaC and NCC, respectively [21,22].

The CUL3 protein, encoded by the *CUL3* gene, also constitutes a component of the ubiquitin E3 ligase complex. CUL3 is a scaffold protein, responsible for recruiting subunits of the complex. Mutations in CUL3, therefore, also impair the ubiquitination of WNK1 and WNK4, thus increasing ENaC and NCCT channel activity [23].

Mutations in the *WNK* genes affect expression of either *WNK1* or *WNK4* kinases, while mutations in *KLHL3* or *CUL3* have an effect on the levels of both *WNK* kinases. This possibly explains the increased phenotypic severity observed in patients harboring *KLHL3/CUL3* mutations [24].

#### 11.4.1.5 Diagnosis

A diagnosis of PHA2 should be suspected in individuals with a presentation of hypertension in adolescence or adulthood and associated hyperkalemia in the laboratory findings, despite normal glomerular function. Other supportive laboratory findings include metabolic acidosis, hyperchloremia, suppressed plasma renin levels, and variable serum aldosterone levels. The presence of a first-degree relative with similar findings is also suggestive of a PHA2 diagnosis; however, the absence of a family history does not preclude the diagnosis.

A PHA2 diagnosis is established either in a proband with the phenotype described above or by the identification of mutations in *WNK1*, *WNK4*, *KLHL3*, or *CUL3* genes through molecular genetic testing approaches. These include either serial single-gene testing approach or the use of a multigene panel. In the former, sequence analysis for pathogenic variants in *CUL3/KLHL3* is performed first, followed by deletion/duplication analysis of the gene in which a mutation is identified. *WNK4* is then analyzed, and if no pathogenic variant is detected, gene-targeted deletion/duplication analysis of *WNK1* is considered. Finally, sequencing of *WNK1* is performed. On the other hand, multigene panel diagnosis involves the use of a panel including all genes of interest. The genes included in a panel differ between laboratories, and some of the genes may not be present, hence the right multigene panel has to be carefully chosen by the clinicians.

#### 11.4.1.6 Management

Hypertension and electrolyte imbalances in PHA2 patients are usually corrected with thiazide diuretics. These drugs inhibit NCCT channel and, therefore, prevent sodium reabsorption in the kidney, thus improving BP levels and any secondary complications of hypertension. To this end, lifestyle changes are also encouraged, including avoiding foods high in salt and potassium. Adequate surveillance is needed to monitor electrolyte levels and BP during treatment.

Relatives of the affected individual are also encouraged to be evaluated either through laboratory investigations to detect electrolyte or BP abnormalities or through genetic testing if a pathogenic variant is known to run in the family. Genetic counseling is also offered, which gives the patients the chance to receive information regarding the condition and its implications, allowing them to make informed medical decisions, as well as personal. This can help with family planning, determining, and explaining the potential risk to the offspring, as well as discussing options such as prenatal testing and DNA banking [10].

### 11.4.2 Liddle's syndrome

#### 11.4.2.1 Case report: a 22-year-old woman with hypertension and hypokalemia was presented

Swetha, a 22-year-old lady, with hypertension and hypokalemia was presented. She was investigated for exclusion of secondary forms of hypertension, and laboratory results revealed low PRA, low serum aldosterone levels, and a metabolic alkalosis. Cardiac work-up was normal. Swetha was placed on spironolactone and atenolol and was recommended to take a high-potassium diet. On a follow-up appointment the BP levels had not been successfully lowered, and the potassium levels had not been corrected. The patient revealed that her sister was identically affected. Although rare, Liddle syndrome was suspected as a result of her hypertension, hypokalemic metabolic alkalosis and suppressed PRA, and serum aldosterone levels, in combination with a positive family history. The diagnosis was confirmed by molecular genetic testing.

#### 11.4.2.2 Definition

Liddle syndrome is an autosomal dominant disorder, characterized by early onset salt-sensitive hypertension, accompanied by hypokalemia, metabolic alkalosis, suppressed PRA, and serum aldosterone levels. It is also commonly associated with a family history of early onset hypertension and, sometimes, sudden death. The hypertension and hypokalemia are the result of increased ENaC activity, leading to enhanced sodium reabsorption and potassium excretion in the distal kidney, which also explains the low plasma renin and aldosterone levels [25]. Other associated symptoms include muscle weakness, polyuria and polydipsia, constipation and heart palpitations, which arise as a result of the hypokalemia. Headaches and dizziness often also arise secondary to the high



BP levels [26]. Liddle syndrome presents with variable degrees of severity, and milder forms are often overlooked. The overall prevalence of Liddle syndrome among hypertensive individuals is unknown [26].

#### 11.4.2.3 Genetics

Liddle syndrome (OMIM: 177200) is a genetically heterogeneous condition, caused by germline mutations in *SCNN1A*, *SCNN1B*, or *SCNN1G* genes. These genes encode for the  $\alpha$ ,  $\beta$ , and  $\gamma$  subunits of the ENaC, respectively. *SCNN1A* maps to chromosome 12p13.31, while *SCNN1B* and *SCNN1G* map to chromosome 16p12.2 [26].

##### 11.4.2.3.1 *SCNN1B* gene

Kindred studies by Liddle et al. as well as Shimkets et al. showed complete linkage of the condition to the gene *SCNN1B* encoding for the  $\beta$  subunit of ENaC channel on chromosome 16. Analysis of the gene demonstrated a premature stop codon (Arg564Ter) which led to truncation of the carboxyl (C) terminus of the encoded protein and loss of the PY motif [27,28]. Mutations in the same carboxyl-terminal domain, including premature termination as well as frameshift mutations were also identified in affected subjects from four different kindreds, confirming that mutations in the *SCNN1B* gene caused Liddle syndrome.

Hansson et al. also identified a missense mutation (Pro616Leu) in *SCNN1B* gene in a different kindred with Liddle syndrome. The presence of this mutation in sodium channels of *Xenopus* oocytes led to increased activation of the channels, demonstrating the significance of the mutation in the pathogenesis of Liddle syndrome [29]. Two bp downstream from the P616L missense mutation described by Hansson et al., Tamura et al. identified a tyr618-to-his mutation, leading to the conclusion that the genetic region between pro616 and try618 has great importance for the regulation of ENaC activity [30].

In a kindred with mild hypertension and low aldosterone levels, a frameshift mutation in the C terminus of *SCNN1B* was also identified. This led to the addition of a cytosine residue at codon 592, leading to a deletion of the last 45 amino acids of the normal protein, including the PY motif (proline-rich target). Sixteen at-risk relatives of the index case were tested for this mutation, and all who were positive for the mutation had mild hypertension, suppressed PRA, and aldosterone levels, while four also had a history of hypokalemia. This demonstrated the phenotypic variability in the severity of the condition and the possibility of Liddle syndrome being underdiagnosed among patients with hypertension [31]. A 32-bp deletion in *SCNN1B* was also identified in a mother and her three sons, leading to a premature stop codon at position 582 [32].

The importance of the PY motif in the regulation of the sodium channel activity, also demonstrated through previously identified mutations, was shown by Inoue et al. in 1997, who identified a missense mutation (Pro615Ser) in the PY motif of the  $\beta$  subunit in a Japanese family [33]. Finally, Furuhashi et al. identified a transversion mutation that led to a pro616-to-arg substitution in the PY motif of the  $\beta$  subunit of ENaC [34]. In total, 24 mutations have been reported in the *SCNN1B* gene [26].

##### 11.4.2.3.2 *SCNN1G* gene

Hansson et al. were the first to discover a nonsense mutation (Trp574Ter) in the  $\gamma$  subunit of the sodium channel in patients with Liddle syndrome. This substitution mutation was also located in the cytoplasmic carboxyl terminus and was found to cause deletion of the last 76 amino acids from the normal protein, also causing a loss of the PY motif. Functional studies in *Xenopus* oocytes with channels expressing the mutation displayed enhanced sodium conductance compared to oocytes with wild-type channels [29]. Six mutations in the *SCNN1G* gene have been reported to date [26].

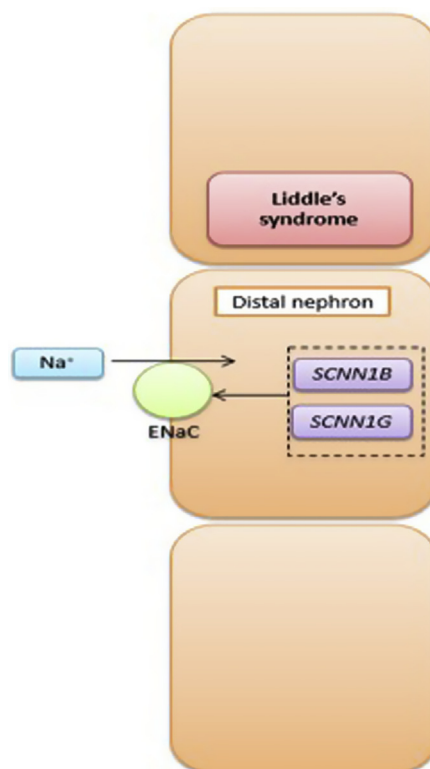
##### 11.4.2.3.3 *SCNN1A* gene

Recently, a missense mutation (Cys479Arg) was identified in the  $\alpha$  subunit of the ENaC channel in a family of Caucasian origin. The mutation was located in the highly conserved extracellular domain of the subunit and disrupted the disulfide bridge, increasing the open conformation of the channel [35]. This is the only mutation in the gene encoding for the  $\alpha$  subunit reported to date [26].

#### 11.4.2.4 Pathophysiology

The ENaC channel is an ENaC, located in the apical membrane of the distal nephron, as well as in the colon, lungs, and exocrine glands. In the kidney, its expression and function are regulated by aldosterone and vasopressin levels, while it plays an important role in sodium homeostasis. ENaC is a heterotrimer consisting of  $\alpha$ ,  $\beta$ , and  $\gamma$  subunits, which are encoded by *SCNN1A*, *SCNN1B*, and *SCNN1G*, respectively. The three subunits share approximately 30% amino acid identity, while they all consist of a short intracellular amino N-terminus, a short





**FIGURE 11.2** Effect of mutations on sodium channel genes in the kidney and the hypertensive phenotype in Liddle's syndrome. Mutations in *SCNN1B* and *SCNN1G* genes increase the density of ENaC channels on the apical membrane of the distal nephron and lead to an increased sodium reabsorption and hypertension. ENaC, Epithelial sodium channel.

carboxy C-terminus, two transmembrane domains, and an extracellular loop. In the carboxy terminus the PY motif constitutes the binding site of Nedd4, a member of the ubiquitin ligase family, which is responsible for the internalization and proteasomal degradation of the sodium channel [26].

Mutations in the *SCNN1B* and *SCNN1G* genes which lead to the removal of the PY motif prevent the internalization and degradation of the sodium channels, leading to an increased expression of ENaC channels in the apical membrane of the distal nephron. This in turn leads to an increase in sodium reabsorption and the hypertensive phenotype of Liddle syndrome (Fig. 11.2). This has been demonstrated through functional studies in *Xenopus* oocytes harboring mutations in these genes, where the activity of the sodium channels with the mutations is significantly higher compared to the activity of the wild-type channels [26,29,34]. However, the mutation identified in the  $\alpha$  subunit of the sodium channel did not affect channel density on the apical membrane of the kidney. Instead, it increased the open conformation of the channel, thus leading to a twofold increase in the inward sodium current [35].

The increase in sodium reabsorption as a result of the enhanced density of ENaC on the apical membrane of the distal nephron and the increase in the channel's open conformation is responsible for the high BP, as well as the low PRA and serum aldosterone levels. Following sodium reabsorption by ENaC, three  $\text{Na}^+$  are actively exchanged for two  $\text{K}^+$  via the  $\text{Na}^+/\text{K}^+$ -ATPase. The potassium ions are then excreted through the apical membrane via different channels. The increase in sodium reabsorption, therefore, leads to increased potassium excretion and results in hypokalemia and metabolic alkalosis [26].

#### 11.4.2.5 Diagnosis

A diagnosis of Liddle's syndrome should be suspected if a patient is presented with resistant early onset hypertension, hypokalemia, low renin and aldosterone levels and metabolic alkalosis, irrespective of a positive family history. The diagnosis is confirmed through molecular genetic testing, which involves sequencing of the *SCNN1A*, *SCNN1B*, and *SCNN1G* genes.

Genetic screening should be performed in all first-degree relatives not only because the disease is inherited in an autosomal dominant fashion but also due to the phenotypic variability of the condition even within the same family [26].

#### 11.4.2.6 Management

Liddle syndrome is treated with the use of  $K^+$ -sparing diuretics amiloride or triamterene, which are ENaC antagonists. Use of these drugs was shown to correct the BP levels as well as the electrolyte abnormalities observed in Liddle syndrome patients. Patients are also encouraged to follow a low salt diet. Monitoring of the patients' potassium levels is appropriate although incidence of hyperkalemia is rare given that renal function and potassium intake are normal. Amiloride constitutes a safe alternative for use during pregnancy and corrects BP, potassium, and renin levels [26].

### 11.4.3 Congenital adrenal hyperplasia

CAH encompasses a group of autosomal recessive disorders which affect steroid hormone synthesis. It affects 1 in 5000 births and as with most autosomal recessive disorders, it is more common among the offspring of consanguineous couples. It is most commonly caused as a result of a 21-hydroxylase deficiency (approximately 95% of cases), and these individuals do not present with hypertension. The second most common cause is 11  $\beta$ -hydroxylase deficiency (OMIM: 202010) representing approximately 5%–8% of the cases. Other forms also exist, including 17  $\alpha$ -hydroxylase deficiency, which is rarer.

#### 11.4.3.1 Case report (1): a 12-year-old boy with hypertension and breast development was presented

Jack, a 12-year-old boy to consanguineous parents, with high BP levels was presented. Laboratory results revealed hyponatremia and hypokalemia, with low plasma renin and serum aldosterone levels. On examination, breast development was observed while he also suffered from precocious puberty with his height and weight being above the 97th percentile. CAH due to 11- $\beta$  hydroxylase deficiency was suspected as a result of his early onset hypertension and hypokalemia associated with breast development and precocious puberty. It was also suspected that virilization led to the wrong sex assignment at birth. Molecular genetic testing revealed mutations in the *CYP11B1* gene confirming the diagnosis as well as a XX karyotype. Glucocorticoid replacement therapy was initiated to correct the BP levels and metabolic abnormalities and surgery was offered for reconstruction of the external genitalia.

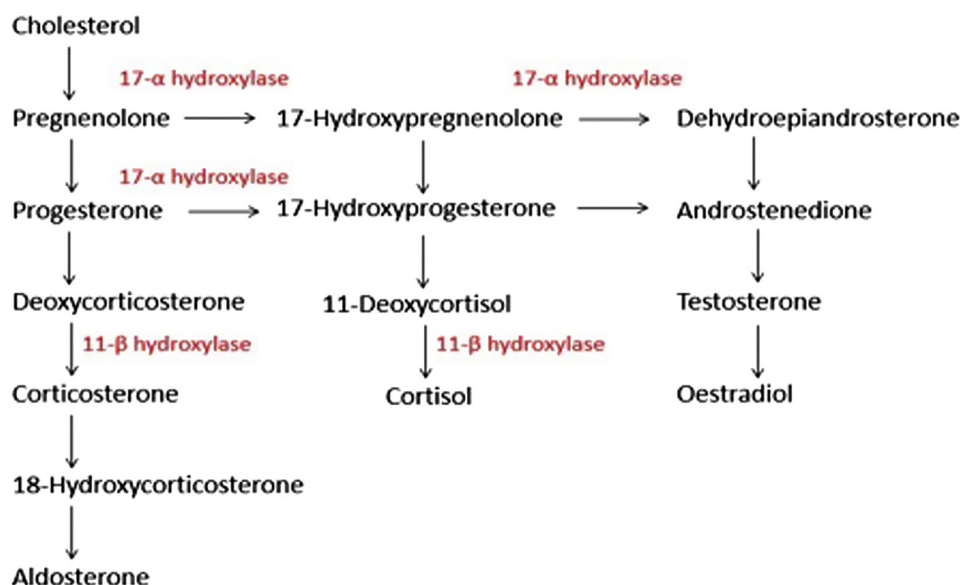
#### 11.4.3.2 Definition of 11 $\beta$ -hydroxylase deficiency

CAH due to 11  $\beta$ -hydroxylase deficiency (OMIM: 202010) is an autosomal recessive disorder of steroid biosynthesis. Its characteristic symptoms include hypertension, androgen excess, and virilization in a background of cortisol deficiency, low aldosterone and high adrenocorticotrophic hormone (ACTH) levels, hypokalemia, and adrenal cortex hypertrophy. The enzyme 11  $\beta$ -hydroxylase is responsible for the production of cortisol and corticosterone. The lack of it, therefore, leads to an accumulation of cortisol and corticosterone precursors and the excessive production of androgens and 11-deoxycorticosterone, which is a mineralocorticoid (Fig. 11.3). The overproduction of these substances explains the virilization and the hypertensive phenotype, respectively.

##### 11.4.3.2.1 Genetics of 11 $\beta$ -hydroxylase deficiency

CAH due to 11  $\beta$ -hydroxylase deficiency is caused primarily by mutations in the *CYP11B1* gene (OMIM: 610613), which is mapped to chromosome 8q24.3 and encodes for the 11  $\beta$ -hydroxylase enzyme.

***CYP11B1* gene** Multiple mutations have been identified in the *CYP11B1* gene amongst patients with 11OHD. White et al. were the first to identify a single point mutation (Arg448His) in exon 8 in 11 of 12 affected individuals of Moroccan Jewish ancestry [36]. Subsequent analysis of the *CYP11B1* gene in other families with 11OHD has revealed over 60 mutations in the *CYP11B1* causing the condition, including premature termination codons, missense mutations, frameshift mutations, splice site changes, insertions, and deletions [13,37–43]. Mutations in the *CYP11B2* gene have also been identified to contribute to 11OHD pathogenesis, due to its close homology to *CYP11B1*. The two genes are located approximately 45 kb apart from each other and are mapped to the same chromosome.



**FIGURE 11.3** Normal pathway of adrenal steroid synthesis. 11- $\beta$  hydroxylase deficiency leads to accumulation of androgens and 11-deoxycorticosterone, a mineralocorticoid; this explains the hypertensive phenotype of this type. 17- $\alpha$  hydroxylase deficiency leads to reduced androgenic steroid and cortisol production. The excess pregnenolone leads to increased deoxycorticosterone production and hence explains the high-blood-pressure levels. Enzymes are listed in red.

#### 11.4.3.2 Pathophysiology of 11 $\beta$ -hydroxylase deficiency

The 11  $\beta$ -hydroxylase enzyme, encoded by the *CYP11B1* gene, is a mitochondrial P450 enzyme located primarily in the zona fasciculata layer of the adrenal cortex. It is responsible for the conversion of 11-deoxycortisol to cortisol and that of 11-deoxycorticosterone to corticosterone. The *CYP11B2* gene encodes for the aldosterone synthase enzyme and is normally expressed in the zona glomerulosa layer of the adrenal glands.

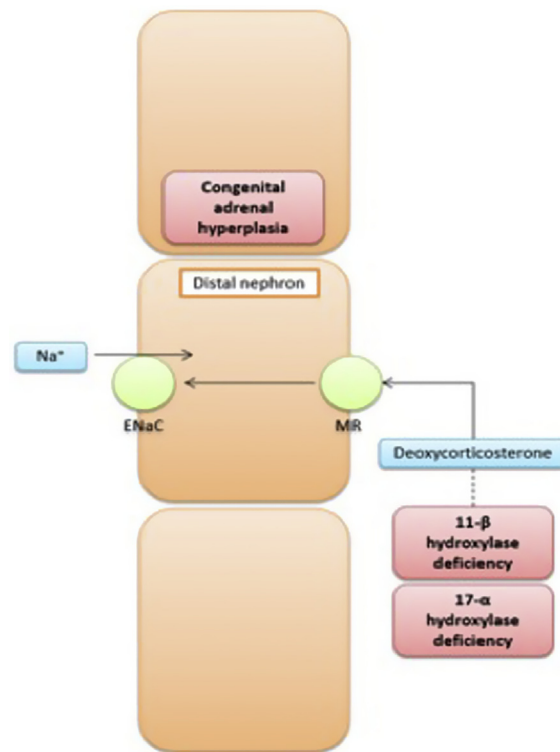
In 11-OHD, mutations in the *CYP11B1* gene lead to the loss of function of the  $\beta$ -hydroxylase enzyme, leading to reduced production of cortisol and corticosterone. Corticosterone is the precursor of aldosterone, and, therefore, the low levels of corticosterone explain the low aldosterone levels in 11-OHD. The reduced levels of cortisol also lead to the activation of the ACTH-dependent feedback mechanism, resulting in the overproduction of steroid precursors, including 11-deoxycortisol and 11-deoxycorticosterone, as well as androgens. Accumulation of 11-deoxycorticosterone, which is a potent mineralocorticoid, leads to increased sodium reabsorption despite low serum aldosterone levels, resulting in the hypertensive phenotype of the condition (Fig. 11.4). Overproduction of androgens is responsible for the symptoms of androgenic excess including virilization, ambiguous external genitalia in females, and a precocious puberty in both sexes, which can result in short stature in adulthood [44].

#### 11.4.3.3 Case report (2): a 20-year-old man with hypertension and ambiguous external genitalia was presented

Jason, a 20-year-old man born of consanguineous parents, was diagnosed with hypertension at the age of 10. He was considered to suffer from early essential hypertension and was prescribed antihypertensive treatment, with no success at controlling BP levels. Laboratory tests revealed a hypokalemic metabolic alkalosis, and low plasma renin and serum aldosterone levels. An ACTH-stimulation test was ordered and demonstrated high deoxycorticosterone, corticosterone, and progesterone levels, and abnormally low cortisol levels. On examination the patient also had a micropenis and a bifid scrotum. A diagnosis of CAH due to 17- $\alpha$  hydroxylase deficiency was suspected. The diagnosis was confirmed through molecular genetic testing. Jason was given glucocorticoid replacement which treated his hypertension and corrected the electrolyte abnormalities.

#### 11.4.3.4 Definition of 17 $\alpha$ -hydroxylase deficiency

CAH due to 17- $\alpha$  hydroxylase deficiency (OMIM: 202110) is a condition of impaired steroid production. It is characterized by hypertension and hypokalemic alkalosis, associated with low renin and aldosterone levels and hypovirilization. The enzyme 17- $\alpha$  hydroxylase is responsible for the hydroxylation of pregnenolone or



**FIGURE 11.4** Effect of enzymes affecting steroid hormone synthesis on kidney channels explains the hypertensive phenotype in CAH 11-OHD and 17-OHD. Deficiency in 11- $\beta$  hydroxylase and 17- $\alpha$  hydroxylase enzymes lead to an accumulation of deoxycorticosterone, a potent mineralocorticoid. This leads to sodium retention in the kidneys and hypertension. ENaC, Epithelial sodium channel; MR, mineralocorticoid receptor.

progesterone and the 17,20 desmolase reaction of 17-hydroxypregnenolone (Fig. 11.3), leading to reduced cortisol and androgenic steroid production. The excess pregnenolone leads to increased deoxycorticosterone production and hence, similar to 11-OHD, hypertension is caused due to an increased mineralocorticoid production (Fig. 11.4).

#### 11.4.3.5 Genetics of 17 $\alpha$ -hydroxylase deficiency

17  $\alpha$ -hydroxylase deficiency is an autosomal recessive condition caused by mutations in the *CYP17* gene, which is mapped to chromosome 10q24.32 and encodes the steroid 17- $\alpha$  hydroxylase enzyme. Mutations in this gene were first discovered by Kagimoto et al. in 1988 who identified a 4-base duplication in exon 8 of the *CYP17A1* gene [45]. This mutation led to loss of enzymatic activity. Since then, multiple mutations in the gene have been identified including missense, nonsense, splice site, deletions, and insertions and leading to either partial or complete loss of enzymatic function [46–49].

#### 11.4.3.6 Pathophysiology of 17 $\alpha$ -hydroxylase deficiency

Mutations in *CYP17A1* gene lead to the loss of the 17- $\alpha$  hydroxylase enzymatic activity. This enzyme belongs in the mitochondrial cytochrome P450 enzyme family and has two distinct actions: hydroxylation within the cortisol synthesis pathway and 17,20 desmolase activity for sex-hormone production. Therefore two forms of the disease are present: complete 17- $\alpha$  hydroxylase/17,20 lyase deficiency, and partial 17- $\alpha$  hydroxylase/17,20 lyase deficiency, which is typically less severe. This enzymatic deficiency leads to reduced cortisol and androgenic steroid production and an accumulation of corticosterone and deoxycorticosterone, which like 11-OHD leads to hypertension (Fig. 11.4). The increased sodium retention caused by the excess deoxycorticosterone leads to an inhibition in the RAAS system, which explains the low renin and serum aldosterone levels. Finally, the low androgen levels lead to the phenotypic presentation of incomplete masculinization and pseudohermaphroditism in males and the primary amenorrhea in females [50].

#### 11.4.4 Diagnosis of congenital adrenal hyperplasia

Diagnosis of CAH should be suspected by clinicians in infants with ambiguous genitalia. The ACTH-stimulation test can be used to measure precursor-product ratios. In 11-OHD the test reveals high 11-deoxycortisol and 11-deoxycorticosterone levels. Patients affected with 17-OHD will also have high levels of 11-deoxycorticosterone, as well as high corticosterone, 17-deoxysteroids and progesterone levels. High follicle-stimulating hormone and luteinizing hormone levels, and low levels of urinary 17-hydroxysteroid-derived compounds are also observed [51]. Molecular genetic testing to screen for mutations in the *CYP11B1* and *CYP17A1* genes are required for confirmation of the diagnosis and sex assignment in 11-OHD and 17-OHD, respectively.

#### 11.4.5 Management of congenital adrenal hyperplasia

The primary aim of CAH treatment involves glucocorticoid replacement in order to suppress ACTH overproduction. Therefore symptomatic management of CAH 11  $\beta$ -hydroxylase deficiency involves the administration of glucocorticoids including dexamethasone and hydrocortisone. Hydrocortisone is usually preferred as it is short-acting and can be administered in pulses to mimic the body's natural cortisol production. Monitoring of the hormonal levels is essential to assess treatment efficacy and to titrate the medication dose accordingly. In patients with 17-OHD, administration of an antihypertensive drug such as a mineralocorticoid antagonist or a calcium channel blocker may also be considered, if glucocorticoid replacement therapy is not sufficient for BP control. Surgical treatment is also available for reconstruction of ambiguous genitalia.

### 11.5 Familial hyperaldosteronism; type 1—glucocorticoid remediable aldosteronism

#### 11.5.1 Case report: a 18-year-old male student with hypertension and hyperaldosteronemia was presented

Jack, an 18-year-old student, with early onset hypertension was presented. On questioning, she revealed that his mum was also diagnosed with hypertension at a young age. Investigations for monogenic forms of hypertension were initiated. Laboratory tests showed high aldosterone levels with suppressed PRA and a metabolic alkalosis. Further investigations revealed normal ACTH levels but elevated plasma 18-oxocortisol and 18-hydroxycortisol. A diagnosis of glucocorticoid-remediable aldosteronism was highly suspected, and a computerised topography (CT) of the abdomen was immediately ordered, which showed a left adrenal hyperplasia. Molecular genetic testing was ordered to confirm the diagnosis, and it came back positive for the chimeric gene causing the condition. Glucocorticoid-replacement therapy with dexamethasone was initiated which corrected Jack's BP levels, as well as the biochemical abnormalities.

#### 11.5.2 Definition

GRA (OMIM: 103900) is the most common form of monogenic hypertension, while it represents 1% of the cases of primary aldosteronism. It is characterized by dexamethasone-suppressible primary hyperaldosteronism and early onset hypertension, which varies in severity among affected individuals. Despite the phenotypic variability in BP levels, all affected individuals demonstrate structural changes in the ventricles of the heart, putting them at risk of cardiovascular complications, while they are also at increased risk of cerebrovascular complications, particularly hemorrhagic strokes. Despite the high aldosterone levels, most individuals are normokalemic, but hypokalemia can be caused by the administration of potassium-wasting diuretics [52]. Laboratory tests also show abnormal plasma renin levels, metabolic alkalosis and elevated 18-oxocortisol and 18-hydroxycortisol levels, while most patients with GRA have either a unilateral aldosterone-producing adrenal adenoma or bilateral idiopathic hyperplasia of the adrenal glands [53]. The condition is caused by the unequal crossing over between the *CYP11B1* and *CYP11B2* genes, which are responsible for the production of the 11- $\beta$  hydroxylase and aldosterone synthase enzymes, respectively. This leads to the ectopic expression of aldosterone synthase in the zona fasciculata layer of the adrenal glands, which is controlled by ACTH secretion, hence explaining the high aldosterone and low plasma renin levels, as well as the hypertensive phenotype [54].



### 11.5.3 Genetics

GRA is an autosomal dominant disorder caused by mutations in the *CYP11B1* gene. It was first identified by Lifton et al. in 1992, who discovered a chimeric gene of the anti-Lepore type, where the 5-prime regulatory sequences of *CYP11B1* (11  $\beta$ -hydroxylase) were fused to the coding region of *CYP11B2* (aldosterone synthase), leading to ectopic expression of aldosterone synthase in the zona fasciculata. The *CYP11B1* gene encodes the 11  $\beta$  hydroxylase enzyme, responsible for the conversion of 11-deoxycortisol to cortisol and 11-deoxycorticosterone to corticosterone in the zona fasciculata of the adrenal cortex (Fig. 11.3). The *CYP11B2* gene encodes for aldosterone synthase, which is normally expressed in the zona glomerulosa and catalyzes the conversion of 11-deoxycorticosterone to corticosterone and corticosterone to 18-hydroxycorticosterone and of 18-hydroxycorticosterone to aldosterone. The two genes show close homology, and they are both mapped to chromosome 8q21 (OMIM, 2009). Affected individuals have two normal copies of both *CYP11B1* and *CYP11B2* as well as the chimeric duplication. DNA-sequencing analysis has also indicated that there is variability in the crossing over and fusion sites among affected individuals, implying that the mutations are not inherited from a single ancestor but, instead, arise independently in each pedigree [55].

### 11.5.4 Pathophysiology

The aldosterone synthase enzyme is normally synthesized in the zona glomerulosa layer of the adrenal cortex as a result of the zonal-specific expression of the *CYP11B2* gene. Consequently, the production of aldosterone is limited to the zona glomerulosa, controlled by the renin–angiotensin system and potassium balance. Similarly, under normal physiological conditions, cortisol production takes place only in the zona fasciculata layer of the adrenal cortex through the hydroxylation of 11-deoxycortisol by 11  $\beta$ -hydroxylase. In GRA the chimeric gene duplication results in the ectopic production of aldosterone synthase and consequently aldosterone synthesis takes place in the fasciculata instead. Aldosterone synthesis in the zona fasciculata is controlled by ACTH secretion, resulting in a circadian pattern in its secretion mimicking that of cortisol. As ACTH is not responsive to high sodium states, hyperaldosteronism occurs leading to a high-volume state and hypertension (Fig. 11.5). The excess aldosterone levels negatively feedback to plasma renin, explaining the abnormal levels of PRA [52].

The exposure of cortisol and corticosterone to aldosterone synthase in the zona fasciculata also leads to the production of 18-oxocortisol and 18-hydroxycortisol, which is another main feature of the condition. Both hybrid steroids have been assessed as potential diagnostic markers for differential diagnosis of primary aldosteronism, and although levels are increased in GRA patients, there is limited data to date to support their application for specific diagnosis [56].

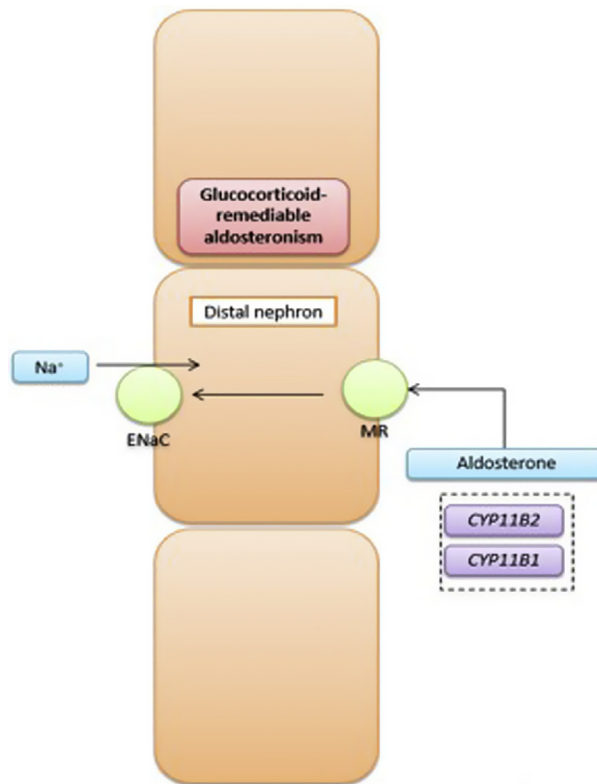
### 11.5.5 Diagnosis

A diagnosis of GRA should be suspected in children presenting with early onset hypertension and a positive family history of early onset hypertension or a hemorrhagic stroke. Patients should also have low-plasma-renin levels, hyperaldosteronemia and elevated 18-hydroxycortisol, 18-hydroxycorticosterone, and 18-oxocortisol levels. However, the diagnosis is confirmed through molecular genetic testing investigating the presence of the chimeric gene [54]. Lifton et al. had originally suggested a Southern blot technique; however, nowadays a faster PCR-based method is used. If access to genetic testing is limited, an alternative involves treating the patient with dexamethasone for 4 days investigating for suppression of aldosterone levels. However, it has been shown that aldosterone levels can be suppressed by dexamethasone in patients lacking the chimeric gene and, therefore, this method may be misleading. Once a patient is diagnosed with GRA, all first-degree relatives should be screened for the disease [52].

### 11.5.6 Management

GRA is effectively managed by exogenous glucocorticoid administration. This results in the suppression of ACTH production, thus suppressing aldosterone synthesis and reversing the clinical and biochemical features of the disease. The exogenous glucocorticoid should have a long biologic half-life, such as dexamethasone or prednisone, while dosing should take place at night to avoid the early morning ACTH surges. Monitoring of the





**FIGURE 11.5** The effect of the chimeric mutation *CYP11B1/CYP11B2* on kidney channels explains the hypertensive phenotype in **GRA**. Unequal cross over between *CYP11B1* and *CYP11B2* leads to the ectopic expression of aldosterone synthase in the zona fasciculata of the adrenal cortex. This leads to high aldosterone levels, increased sodium reabsorption, and hypertension. *ENaC*, Epithelial sodium channel; *MR*, mineralocorticoid receptor.

plasma renin levels and plasma aldosterone concentrations should be used to guide the treatment. It should be noted that 18-oxosteroid and even aldosterone levels may remain high despite achieving BP target levels [52].

In case the first-line treatment is ineffective in lowering BP levels, a mineralocorticoid receptor antagonist should be considered, including spironolactone and eplerenone. Eplerenone is the preferred drug of choice as it is a more specific antagonist, leading to fewer side effects. Amiloride or triamterene inhibit the ENaC in the distal nephron, thus inhibiting sodium reabsorption and could be alternatively used for the management of GRA, together with other antihypertensive agents. Finally, all diagnosed patients are advised to follow a low sodium diet to avoid hypokalemia and help with their BP management [52].

## 11.6 Genetic overlap of monogenic and essential hypertension

As the number of individuals with genetic data continues to increase, larger and larger GWAS for BP traits are being undertaken. There are now more than 1000 BP loci described in the literature, and with such large datasets, it is possible to check for variation in genes causing monogenic forms of hypertension in individuals with essential hypertension. The largest BP GWAS to date included over 1 million individuals and reported 901 BP loci, of which 535 were new. A review of genome-wide significant variants/loci has provided validation of prior genome-wide significant associations for variation in two monogenic BP genes: *CYP17A1* and *PDE3A*. Several associated variants were observed at each locus, both common and low frequency. None of the variants were those associated with monogenic disease as expected. The second approach that has been taken for checking overlap of variation in genes causing monogenic and essential hypertension is testing for enrichment of rare and common variants in selected monogenic genes in the general population. Two recent studies using exome array data found no evidence of enrichment of rare variants across the panel of candidate genes tested [57,58].

However, analyses of exome array data did reveal significant association of low frequency and rare variants in other genes (*RRAS*, *COL21A1*, *DBH*, *NPR1*, and *RBM47*), indicating there are rare variants in the general population associated with hypertension [57,58]. Thus results so far indicate there is minimal overlap of genetic variation in genes and mechanisms causing monogenic forms of hypertension and essential hypertension in the general population.

### 11.7 Clinical implications from genetic studies of hypertension

The identification of mutations causing monogenic forms of hypertension has provided targeted screening for patients permitting an early diagnosis and guidelines for management. With whole genome sequencing (WGS) now being deployed for gene identification of rare inherited diseases following negative panel testing (if no mutations are found in candidate genes), further mutations causing rare forms of hypertension in families are expected. In the United Kingdom the 100,000 Genomes Project (a national program to enable new scientific discovery and medical insights for rare diseases and common cancers) has now completed the sequencing of over 85,000 whole genomes and is providing new gene discoveries. The UK government has recently announced WGS of a further 5 million individuals, this investment into genomic medicine will revolutionize medical practice for diagnosis of inherited disease over the next decade. Similar initiatives are also on-going worldwide, for example, the National Human Genome Research Institute—Genome Sequencing Program in the United States, WGS is being performed on individuals with both common diseases and Mendelian conditions, and in Canada the WGS of the first 56 individuals was reported in 2018 [59]. Data from sequencing efforts are also being provided to researchers in public databases, for example, Gnom-AD, a resource with over 100,000 exome sequences and 15,000 WGS from disease cases and population samples [60]. If we review new knowledge from genetic studies of essential hypertension, GWAS discoveries have validated many known BP therapeutic targets (*PKD2LI*, *SLC12A2*, *CACNA1C*, *CACNB4*, and *CA7*) and identified new targets and indicate therapeutics for possible repurposing. With the increase of BP loci, improved genetic risk scores (GRSs) have been created including variants at the reported 901 BP loci. Results indicate significant associations, with a 10.4 mmHg higher sex-adjusted mean SBP comparing the top and bottom quintiles of the GRS distribution ( $P < 1 \times 10^{-300}$ ). The same GRS is associated with a threefold increased risk of hypertension comparing the top and bottom quintiles of the GRS distribution ( $P < 1 \times 10^{-300}$ ). These data provide further support for genotyping BP variants in young individuals, so individuals at risk of hypertension are identified earlier, and there is the opportunity for lifestyle management of BP [9].

### 11.8 Future perspectives

The deployment of genotyping and dense imputation panels in samples from large-scale national biobanks is facilitating additional gene discoveries across all inherited traits. With the inclusion of up to 500,000 individuals from UK Biobank, we have witnessed a trebling in the number of BP loci discoveries over the past 18 months, and 27% of the estimated heritability of BP can now be explained [9]. The results from GWAS also provide a wealth of new information on BP regulation processes and potential therapeutic targets for treating hypertension and cardiovascular disease. There is however a bottleneck in moving from a GWAS signal to identifying the causal gene and mechanisms, but over the next few years with continual development of new statistical methods, forthcoming WGS datasets and shared bioinformatic resources (e.g., ENCODE, Roadmap Epigenomics, and Fantom5) quicker progress is expected [61]. Machine learning and artificial intelligence are also being deployed for analysis of WGS, electronic health record data and additional “omic” datasets, this is an area of research ripe for further exploitation. A proof of concept study using machine learning has recently been published for abdominal aortic aneurysm providing identification of disease loci for this trait, and GWAS have not been successful. This study also introduced information for health management [62]. Over the next few years, we expect an enhanced understanding of genetic factors for hypertension, and as further genetic variants are discovered including rare variants with larger effect sizes and incorporated into GRSs, we expect improved prediction of an individual’s susceptibility to hypertension, this should enable healthcare practitioners to start considering tailored management plans for the prevention and management of hypertension.

## Acknowledgments

We acknowledge the support of the National Institute for Health Research Biomedical Research Centre at Barts, Queen Mary University of London.

## References

- [1] Bloch MJ. Worldwide prevalence of hypertension exceeds 1.3 billion. *J Am Soc Hypertens* 2016;10(10):753–4. Available from: <https://doi.org/10.1016/j.jash.2016.08.006>.
- [2] Makridakis S, DiNicolantonio JJ. Hypertension: empirical evidence and implications in 2014. *Open Heart* 2014;1(1):e000048. Available from: <https://doi.org/10.1136/openhrt-2014-000048>.
- [3] Rossi GP, et al. Genetic screening in arterial hypertension. *Nat Rev Endocrinol* 2017;13(5):289–98. Available from: <https://doi.org/10.1038/nrendo.2016.196>.
- [4] World Health Organization. *A global brief on hypertension*, WHO. World Health Organization; 2013. Available at: [http://www.who.int/cardiovascular\\_diseases/publications/global\\_brief\\_hypertension/en/](http://www.who.int/cardiovascular_diseases/publications/global_brief_hypertension/en/) [accessed 04.11.18].
- [5] Oparil S, et al. Hypertension. *Nature reviews disease primers*, 4. Nature Publishing Group; 2018. p. 18014. <https://doi.org/10.1038/nrdp.2018.14>.
- [6] Brettles J. Essential hypertension. Available at: <https://bestpractice.bmj.com/topics/en-gb/26>; 2018 [accessed 10.09.18].
- [7] Pazoki R, et al. Genetic predisposition to high blood pressure and lifestyle factors. *Circulation* 2018;137(7):653–61. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.117.030898>.
- [8] Hottenga J-J, et al. Heritability and stability of resting blood pressure. *Twin Res Hum Genet* 2005;8(5):499–508. Available from: <https://doi.org/10.1375/twin.8.5.499>.
- [9] Evangelou E, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet* 2018;50(10):1412–25. Available from: <https://doi.org/10.1038/s41588-018-0205-x>.
- [10] Ellison DH. *Pseudohypoaldosteronism Type II*, *GeneReviews*®. Seattle: University of Washington; 1993. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22073419> [accessed 02.11.18].
- [11] Boyden LM, et al. Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature* 2012;482(7383):98–102. Available from: <https://doi.org/10.1038/nature10814>.
- [12] Wilson FH, et al. Human hypertension caused by mutations in WNK kinases. *Science* 2001;293(5532):1107–12. Available from: <https://doi.org/10.1126/science.1062844>.
- [13] Louis-Dit-Picard H, et al. KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. *Nat Genet* 2012;44(4):456–60. Available from: <https://doi.org/10.1038/ng.2218>.
- [14] Medicinal Chemistry of Anticancer Drugs. Protein kinases. Available at: <https://www.sciencedirect.com/topics/neuroscience/protein-kinases>; 2008 [accessed 02.11.18].
- [15] Cheng C-J, Huang C-L. Activation of PI3-kinase stimulates endocytosis of ROMK via Akt1/SGK1-dependent phosphorylation of WNK1. *J Am Soc Nephrol* 2011;22(3):460–71. Available from: <https://doi.org/10.1681/ASN.2010060681>.
- [16] Rodan AR, Jenny A. WNK kinases in development and disease. *Curr Top Dev Biol* 2017;123:1–47. Available from: <https://doi.org/10.1016/bs.ctdb.2016.08.004>.
- [17] Hoorn EJ, Ellison DH. WNK kinases and the kidney. *Exp Cell Res* 2012;318(9):1020–6. Available from: <https://doi.org/10.1016/j.yexcr.2012.02.029>.
- [18] Yang C-L, et al. WNK kinases regulate thiazide-sensitive Na-Cl cotransport. *J Clin Invest* 2003;111(7):1039–45. Available from: <https://doi.org/10.1172/JCI17443>.
- [19] Naray-Fejes-Toth A, Snyder PM, Fejes-Toth G. The kidney-specific WNK1 isoform is induced by aldosterone and stimulates epithelial sodium channel-mediated Na<sup>+</sup> transport. *Proc Natl Acad Sci USA* 2004;101(50):17434–9. Available from: <https://doi.org/10.1073/pnas.0408146101>.
- [20] Vidal-Petiot E, et al. WNK1-related familial hyperkalemic hypertension results from an increased expression of L-WNK1 specifically in the distal nephron. *Proc Natl Acad Sci USA* 2013;110(35):14366–71. Available from: <https://doi.org/10.1073/pnas.1304230110>.
- [21] Susa K, et al. Impaired degradation of WNK1 and WNK4 kinases causes PHAII in mutant KLHL3 knock-in mice. *Hum Mol Genet* 2014;23(19):5052–60. Available from: <https://doi.org/10.1093/hmg/ddu217>.
- [22] Wu G, Peng J-B. Disease-causing mutations in KLHL3 impair its effect on WNK4 degradation. *FEBS Lett* 2013;587(12):1717–22. Available from: <https://doi.org/10.1016/j.febslet.2013.04.032>.
- [23] Hershko A, Ciechanover A. THE ubiquitin system. *Annu Rev Biochem* 1998;67(1):425–79. Available from: <https://doi.org/10.1146/annurev.biochem.67.1.425>.
- [24] Sohara E, Uchida S. Kelch-like 3/Cullin 3 ubiquitin ligase complex and WNK signaling in salt-sensitive hypertension and electrolyte disorder. *Nephrol Dial Transp* 2016;31(9):1417–24. Available from: <https://doi.org/10.1093/ndt/gfv259>.
- [25] Hechanova LA. Liddle syndrome. Available at: <https://www.msdmanuals.com/en-gb/professional/genitourinary-disorders/renal-transport-abnormalities/liddle-syndrome>; 2018 [accessed 02.11.18].
- [26] Tetti M, et al. Liddle syndrome: review of the literature and description of a new case. *Int J Mol Sci* 2018;19(3):812. Available from: <https://doi.org/10.3390/ijms19030812>.
- [27] Liddle GW. A familial renal disorder simulating primary aldosteronism but with negligible aldosterone secretion. *Trans Assoc Am Physicians* 1963;76:199–213. Available at: <https://ci.nii.ac.jp/naid/10018198924> [accessed 02.11.18].
- [28] Shimkets RA, et al. Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. *Cell* 1994;79(3):407–14. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/7954808> [accessed 02.11.18].
- [29] Hansson JH, et al. Hypertension caused by a truncated epithelial sodium channel  $\gamma$  subunit: genetic heterogeneity of Liddle syndrome. *Nat Genet* 1995;11(1):76–82. Available from: <https://doi.org/10.1038/ng0995-76>.

- [30] Tamura H, et al. Liddle disease caused by a missense mutation of beta subunit of the epithelial sodium channel gene. *J Clin Invest* 1996;97(7):1780–4. Available from: <https://doi.org/10.1172/JCI118606>.
- [31] Findling JW, et al. Liddle's syndrome: prospective genetic screening and suppressed aldosterone secretion in an extended kindred. *J Clin Endocrinol Metab* 1997;82(4):1071–4. Available from: <https://doi.org/10.1210/jcem.82.4.3862>.
- [32] Jeunemaitre X, et al. Genotype-phenotype analysis of a newly discovered family with Liddle's syndrome. *J Hypertens* 1997;15(10):1091–100. Available at: <<http://www.ncbi.nlm.nih.gov/pubmed/9350583>> [accessed 03.11.18].
- [33] Inoue J, Iwaoka T, Tokunaga H, Takamune K, Naomi S, Araki M, Takahama K, Yamaguchi K, Tomita K. A family with Liddle's syndrome caused by a new missense mutation in the  $\beta$  subunit of the epithelial sodium channel. *J. Clin. Endocrinol. Metab.* 1998;83:2210–13. Available from: <https://doi.org/10.1210/jcem.83.6.5030>.
- [34] Furuhashi M, et al. Liddle's syndrome caused by a novel mutation in the proline-rich PY motif of the epithelial sodium channel  $\beta$ -subunit. *J Clin Endocrinol Metab* 2005;90(1):340–4. Available from: <https://doi.org/10.1210/jc.2004-1027>.
- [35] Salih M, et al. A missense mutation in the extracellular domain of  $\alpha$  ENaC causes Liddle syndrome. *J Am Soc Nephrol* 2017;28(11):3291–9. Available from: <https://doi.org/10.1681/ASN.2016111163>.
- [36] White PC, et al. A mutation in CYP11B1 (Arg-448→His) associated with steroid 11 beta-hydroxylase deficiency in Jews of Moroccan origin. *J Clin Invest* 1991;87(5):1664–7. Available from: <https://doi.org/10.1172/JCI115182>.
- [37] Ezquieta B, Luzuriaga C. Neonatal salt-wasting and 11  $\beta$ -hydroxylase deficiency in a child carrying a homozygous deletion hybrid CYP11B2 (aldosterone synthase)-CYP11B1 (11  $\beta$ -hydroxylase). *Clin Genet* 2004;66(3):229–35. Available from: <https://doi.org/10.1111/j.1399-0004.2004.00291.x>.
- [38] Hampf M, et al. Unequal crossing-over between aldosterone synthase and 11 $\beta$ -hydroxylase genes causes congenital adrenal hyperplasia. *J Clin Endocrinol Metab* 2001;86(9):4445–52. Available from: <https://doi.org/10.1210/jcem.86.9.7820>.
- [39] Helmborg A, Ausserer B, Kofler R. Frame shift by insertion of 2 basepairs in codon 394 of CYP11B1 causes congenital adrenal hyperplasia due to steroid 11 beta-hydroxylase deficiency. *J Clin Endocrinol Metab* 1992;75(5):1278–81. Available from: <https://doi.org/10.1210/jcem.75.5.1430088>.
- [40] Joehrer K, et al. CYP11B1 mutations causing non-classic adrenal hyperplasia due to 11-hydroxylase deficiency. *Hum Mol Genet* 1997;6(11):1829–34. Available from: <https://doi.org/10.1093/hmg/6.11.1829>.
- [41] Krone N, et al. Analyzing the functional and structural consequences of two point mutations (P94L and A368D) in the CYP11B1 gene causing congenital adrenal hyperplasia resulting from 11-hydroxylase deficiency. *J Clin Endocrinol Metab* 2006;91(7):2682–8. Available from: <https://doi.org/10.1210/jc.2006-0209>.
- [42] Matallana-Rhoades AM, et al. Congenital adrenal hyperplasia due to 11-beta-hydroxylase deficiency: description of a new mutation, R384X. *Colombia Médica* 2016;47(3):172–5. Available at: <<http://colombiamedica.univalle.edu.co/index.php/comedica/article/view/1722>> [accessed 03.11.18].
- [43] Naiki Y, et al. A nonsense mutation (TGG [Trp116]→TAG [Stop]) in CYP11B1 causes steroid 11 beta-hydroxylase deficiency. *J Clin Endocrinol Metab* 1993;77(6):1677–82. Available from: <https://doi.org/10.1210/jcem.77.6.7903314>.
- [44] Menabò S, et al. Congenital adrenal hyperplasia due to 11-beta-hydroxylase deficiency: functional consequences of four CYP11B1 mutations. *Eur J Hum Genet* 2014;22(5):610–16 <<https://doi.org/10.1038/ejhg.2013.197>>.
- [45] Kagimoto M, et al. Structural characterization of normal and mutant human steroid 17 $\alpha$ -hydroxylase genes: molecular basis of one example of combined 17 $\alpha$ -hydroxylase/17,20 lyase deficiency. *Mol Endocrinol* 1988;2(6):564–70. Available from: <https://doi.org/10.1210/mend-2-6-564>.
- [46] Biglieri EG, Herron MA, Brust N. 17-Hydroxylation deficiency in man. *J Clin Invest* 1966;45(12):1946–54. Available from: <https://doi.org/10.1172/JCI105499>.
- [47] Geller DH, et al. The genetic and functional basis of isolated 17,20-lyase deficiency. *Nat Genet* 1997;17(2):201–5. Available from: <https://doi.org/10.1038/ng1097-201>.
- [48] Martin RM, et al. P450c17 deficiency in Brazilian patients: biochemical diagnosis through progesterone levels confirmed by CYP17 genotyping. *J Clin Endocrinol Metab* 2003;88(12):5739–46. Available from: <https://doi.org/10.1210/jc.2003-030988>.
- [49] Yanase T, Simpson ER, Waterman MR. 17 $\alpha$ -Hydroxylase/17,20-lyase deficiency: from clinical investigation to molecular definition. *Endocr Rev* 1991;12(1):91–108. Available from: <https://doi.org/10.1210/edrv-12-1-91>.
- [50] Hari Krishnan Parthasarathy HK, Macdonald T. Review: aldosterone antagonists. *Br J Diab Vas Dis* 2008;8(5):215–19. Available from: <https://doi.org/10.1177/1474651408094591>.
- [51] Yau M, New M. Congenital adrenal hyperplasia: diagnosis and emergency treatment. *Endotext* 2000;. Available at: <<http://www.ncbi.nlm.nih.gov/pubmed/25905311>> [accessed 03.11.18].
- [52] Halperin F, Dluhy RG. Glucocorticoid-remediable aldosteronism. *Endocrinol Metab Clin North Am* 2011;40(2):333–41. Available from: <https://doi.org/10.1016/j.ecl.2011.01.012>.
- [53] McMahon GT, Dluhy RG. Glucocorticoid-remediable aldosteronism. *Cardiol Rev* 2004;12(1):44–8. Available from: <https://doi.org/10.1097/01.crd.0000096417.42861.ce>.
- [54] Shahrrava A, et al. A case of glucocorticoid remediable aldosteronism and thoracoabdominal aneurysms. *Case Rep Endocrinol* 2016;2016:1–4 2017571. <<https://doi.org/10.1155/2016/2017571>>.
- [55] Dluhy RG, Williams GH. Glucocorticoid-remediable aldosteronism. *Cardiovasc Res* 1996;31(6):870–2. Available at: <<http://www.ncbi.nlm.nih.gov/pubmed/8759241>> [accessed 03.11.18].
- [56] Lenders JWM, et al. 18-Oxocortisol and 18-hydroxycortisol: is there clinical utility of these steroids? *Eur J Endocrinol* 2018;178(1):R1–9. Available from: <https://doi.org/10.1530/EJE-17-0563>.
- [57] Liu C, et al. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nat Genet* 2016;48(10):1162–70. Available from: <https://doi.org/10.1038/ng.3660>.
- [58] Surendran P, et al. Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nat Genet* 2016;48(10):1151–61. Available from: <https://doi.org/10.1038/ng.3654>.

- [59] Reuter MS, et al. The Personal Genome Project Canada: findings from whole genome sequences of the inaugural 56 participants. *Can Med Assoc J* 2018;190(5):E126–36. Available from: <https://doi.org/10.1503/cmaj.171151>.
- [60] Lek M, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536(7616):285–91. Available from: <https://doi.org/10.1038/nature19057>.
- [61] Cannon ME, Mohlke KL. Deciphering the emerging complexities of molecular mechanisms at GWAS loci. *Am J Hum Genet* 2018;103(5):637–53. Available from: <https://doi.org/10.1016/j.ajhg.2018.10.001>.
- [62] Li J, et al. Decoding the genomics of abdominal aortic aneurysm. *Cell* 2018;174(6):1361–1372.e10. Available from: <https://doi.org/10.1016/j.cell.2018.07.021>.





## Molecular basis of stroke

*Yoshiji Yamada*

Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University,  
Tsu, Mie, Japan

### 12.1 Introduction

Stroke is a complex multifactorial disorder that is thought to result from an interaction between a person's genetic background and various environmental factors. It is a common and serious condition, with about 795,000 individuals experiencing a new or recurrent stroke and nearly 150,000 deaths from stroke-related causes in 2008 in the United States. The prevalence of stroke in the United States is 7 million. Of all such events, 87% are ischemic stroke, 10% are intracerebral hemorrhage, and 3% are subarachnoid hemorrhage [1] (Fig. 12.1). Despite recent advances in acute stroke therapy, stroke remains the leading cause of severe disability and the third leading cause of death, after heart disease and cancer, in Western countries [2]. The identification of biomarkers of stroke risk is important both for risk prediction and for intervention to avert future events.

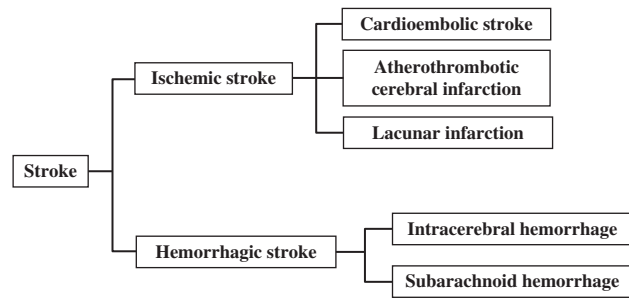
### 12.2 Genetics of stroke

Ischemic and hemorrhagic stroke may have both shared and different determinants, although the genetic variants that influence these clinical conditions are probably different. Studies with twins, siblings, and families have provided substantial evidence for heritability of common forms of stroke [3], but the genetic determinants remain largely unknown. Specific mutations in several monogenic stroke disorders have been identified [4]. Although these observations provide insight into the pathophysiological processes of stroke, these mutations are rare and do not contribute substantially to stroke risk in the general population.

A family history of stroke is regarded as an important risk factor for this disease [5]. A positive family history might be the result of shared genes, a shared environment, or both. Despite the identification of rare Mendelian stroke syndromes in humans [6,7], many candidate gene association studies for common forms of stroke have produced few consistent results and data on the genetic epidemiology of stroke are conflicting [7,8]. The incidence of ischemic stroke, intracerebral hemorrhage, or subarachnoid hemorrhage differs among ethnic groups, which may be attributable to differences in the distribution and frequency of genetic polymorphisms as well as in environmental factors such as diet, exercise, and other lifestyle aspects. Given that some gene polymorphisms characteristic of specific ethnic groups may be related to stroke, it is necessary to examine the relations of gene polymorphisms to stroke in each ethnic group.

### 12.3 Single-gene disorders associated with stroke

Several conditions in which stroke occurs are inherited in a classical Mendelian pattern as autosomal dominant, autosomal recessive, or X-linked disorders [7,9,10]. In most of these conditions, stroke is just one component



**FIGURE 12.1** Classification of stroke. There are two types of stroke; ischemic stroke and hemorrhagic stroke. Ischemic stroke includes atherothrombotic cerebral infarction, cardioembolic stroke, and lacunar infarction. Hemorrhagic stroke includes intracerebral hemorrhage and subarachnoid hemorrhage.

of the disease phenotype, but in others, it is the prominent or sole clinical manifestation [10]. Studies of some Mendelian forms of stroke have identified the genes responsible [11,12].

Cerebral arteriopathy autosomal dominant with subcortical infarcts and leukoencephalopathy (CADASIL), an autosomal dominant form of stroke, has been well characterized genetically and shown to be attributable to mutation of the notch 3 gene (*NOTCH3*). Linkage analysis thus mapped the responsible gene to a defined region of human chromosome 19p13 [13,14], the gene was then isolated by positional cloning, and the mutation was identified and its functional role confirmed [15,16]. A striking feature of CADASIL is that disease severity is highly variable even within families. Recent data suggest the importance of both genetic and environmental modifying factors. Confluent white matter hyperintensities (WMH) on magnetic resonance imaging (MRI) are a key feature in CADASIL. A family study measuring WMH volume within a CADASIL population demonstrated a heritability of 63%, suggesting that other genes interact with *NOTCH3* to modulate the disease phenotype [17].

Genes causing rare monogenic forms of small-vessel disease have recently been identified. Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy causes lacunar stroke and early onset vascular dementia. Individuals also typically have alopecia beginning in the second decade, and spondylosis in the second or third decade. It has been shown to result from mutations in the HtrA serine peptidase 1 gene (*HTRA1*) which is involved in transforming growth factor, beta 1 signaling [18].

Two deadly forms of inherited intracerebral hemorrhage have been described in the Dutch and Icelandic populations [8]. Hereditary cerebral hemorrhage with amyloidosis—Dutch type (HCHWA-D) is due to a mutation in the amyloid beta precursor protein gene (*APP*) [19]. The Icelandic form of this condition (HCHWA-I) is due to mutations in the gene coding for cystatin C (*CST3*), a serine protease inhibitor [20,21]. These disorders are characterized by the development of cerebral hemorrhage at an age of 40–50 years for HCHWA-D and 20–30 years for HCHWA-I. Both are associated with amyloid deposition in cortical and leptomeningeal arterioles [8]. Mutations of the integral membrane protein 2B gene (*ITM2B*) have also been shown to result in autosomal dominant amyloid angiopathies, which lead to cerebral hemorrhage, vascular dementia, or both [22]. The *KRIT1*, ankyrin repeat containing gene (*KRIT1*), has also been identified as one of the genes responsible for cavernous angiomas [23].

Autosomal dominant retinal vasculopathy with cerebral leukodystrophy is a microvascular endotheliopathy presenting with visual loss, stroke, and dementia, with onset in the middle age. C-terminal frameshift mutations in the three prime repair exonuclease 1 gene (*TREX1*), which is ubiquitously expressed in the mammalian cells, were identified [24]. These truncated proteins retain exonuclease activity but lose normal perinuclear localization [11].

Mutations in the collagen, type IV, alpha 1 gene (*COL4A1*) affecting glycine residues that are in close proximity in exons 24 and 25 within the triple helix domain of the protein cause a syndrome of hereditary angiopathy with nephropathy, aneurysms, and muscle cramps (HANAC). Patients with HANAC have disruption of the basement membrane systemically. The nephropathy causes hematuria and renal cysts. The muscle cramps occur with or without elevations in serum creatine kinase. Patients have a characteristic retinal arteriopathy. There has been further characterization of the neurovascular phenotype in a series of families [25]. Infarctions not related to cardiac or large vessel pathology occur at an early age. Patients may be predisposed to posttraumatic hemorrhage. Intracranial aneurysms are characteristically localized to various levels of the carotid siphon [26].

Another Mendelian condition associated with stroke is mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes (MELAS), a genetically heterogeneous mitochondrial disorder with a variable clinical phenotype. It is accompanied by features of central nervous system involvement, including seizures, hemiparesis, hemianopsia, cortical blindness, and episodic vomiting [27,28]. This syndrome has been attributed to single

nucleotide mutations in mitochondrial DNA. The mutations are usually, but not exclusively, missense and lie within the tRNA<sup>Leu(UUR)</sup> gene, with an A→G transition at position 3243 [29] and a T→C transition at position 3271 [30] being most frequently reported. Individuals who have inherited one [31,32] or both [33] of these mutations have a greater predisposition to stroke [12].

Ischemic stroke is occasionally attributable to an underlying connective tissue disorder that results in arterial dissection [12]. In Marfan syndrome, extension of aortic dissection into the common carotid artery can occur and result in stroke [34]. Defects in collagen synthesis in Ehlers–Danlos syndrome type IV can predispose affected individuals to spontaneous dissection of the extracranial carotid and vertebral arteries [35]. Fabry disease is an X-linked disorder caused by a deficiency of alpha-galactosidase and is associated with a high risk of both stroke and coronary heart disease [36].

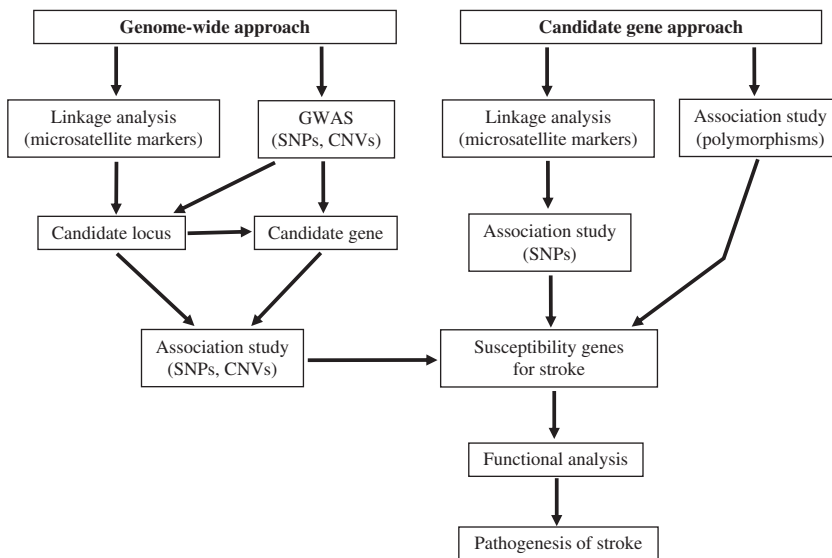
Identification of the genes responsible for Mendelian forms of stroke by reverse genetics has provided new insights into the pathophysiology of stroke [10]. These observations constitute the basis for clinically useful molecular diagnostic tests. Despite their low prevalence, monogenic conditions should always be considered in young patients suffering stroke or in patients of any age with no evidence of vascular risk factors, especially when there is a family history. Indeed, the risk of stroke both in individuals known to have the mutated gene and in their relatives is high. For example, in the case of an autosomal dominant disorder with complete penetrance, all persons who carry the mutated gene will have a stroke, as will half of their first-degree relatives [10].

## 12.4 Genetics of common forms of stroke

The etiology of common forms of stroke is multifactorial and includes both genetic and environmental factors. Studies with families have estimated that the relative risk of stroke in a first-degree relative of an individual who has a stroke is between 1.5 and 2.5. Such a risk is low at the individual level and may not have practical clinical implications. However, this slight increase in the risk of stroke is important at the population level, because of the high incidence of stroke [10]. Identification of genetic variants that contribute to the increased risk of stroke is therefore clinically important.

Common forms of stroke are heterogeneous and most likely results in part from the additive or multiplicative effects of a wide spectrum of pathogenic alleles, each of which confers a small degree of risk. Some of these alleles may predispose individuals to specific types or subtypes of stroke by affecting certain intermediate factors that either lead to stroke, such as the intimal-medial thickness of the carotid artery, or have a direct independent effect on the risk of stroke. In addition, gene variants may also modulate the severity of stroke [10].

In spite of the large number of studies that have identified genes or polymorphisms associated with stroke, only a small number of the findings of these studies have been confirmed by independent replication or in other ethnic groups. One reason for this inconsistency is that many studies have combined ischemic and hemorrhagic stroke and it is unlikely that these different pathological conditions are under the same genetic influences [37]. Another reason is that, although ischemic stroke is a highly complex trait, few studies have assessed subtypes of ischemic stroke or have had sufficient statistical power to do so [37]. Many studies have thus analyzed atherothrombotic cerebral infarction and cardioembolic stroke collectively as ischemic stroke; the former results from the development of atherosclerotic stenosis in carotid or vertebral arteries, whereas the latter is attributable to the obstruction of cerebral arteries by thrombi that are generated in the cardiac atrium or ventricle as a result of arrhythmia such as atrial fibrillation or of valvular or ischemic heart disease. However, it can be argued that atherosclerosis is also responsible for most cardioembolic strokes, given that many such events are the consequence of thrombus formation on the damaged endocardial surface in acute myocardial infarction or within a ventricular aneurysm caused by damage to cardiac muscle secondary to previous myocardial infarction [38]. Most cardioembolic strokes used to be caused by atrial fibrillation secondary to mitral stenosis associated with rheumatic heart disease, but the incidence of childhood rheumatic fever has decreased markedly in the era of penicillin [39,40]. Most cases of atrial fibrillation are now caused by cardiac damage secondary to coronary heart disease. A substantial proportion of cardioembolic strokes is therefore related to atherosclerosis of coronary arteries [38]. The etiology of intracardiac thrombi is diverse, however, including lone atrial fibrillation and other arrhythmias, valvular heart disease, cardiomyopathies, as well as coronary heart disease. Atherothrombotic cerebral infarction and cardioembolic stroke are thus different disorders. Given that the effects of gene polymorphisms or haplotypes on the development of common forms of stroke are likely to be small, it is necessary to examine these disorders separately in order to identify associated genetic variants.



**FIGURE 12.2** Strategies for identifying susceptibility genes for stroke. There are two basic strategies for identifying genes that influence common diseases or other complex traits, the genome-wide approach and the candidate gene approach, both of which rely on linkage analyses and association studies. In the GWAS, SNPs or CNVs distributed throughout the entire genome are used to identify genomic regions that harbor genes that influence the trait of interest with a detectable effect size. The candidate gene approach involves the direct examination of whether an individual gene or genes might contribute to the trait of interest. CNVs, Copy number variations; GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms.

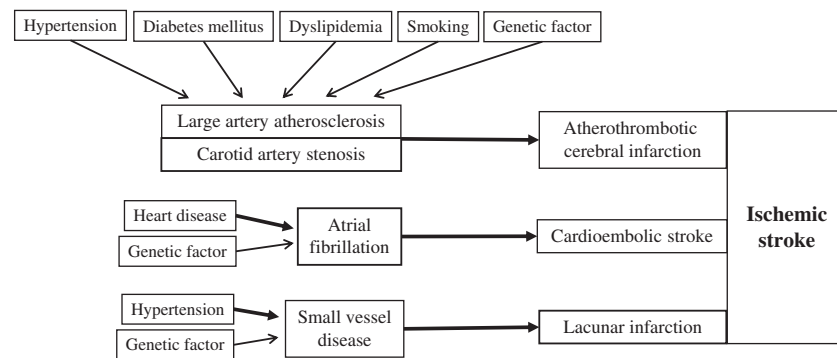
### 12.4.1 Strategies for genetic analysis of stroke

There are two basic strategies for identifying genes that influence the predisposition to stroke: linkage analyses and association studies (Fig. 12.2). Linkage analysis involves the proposition of a model to account for the pattern of inheritance of a phenotype observed in a pedigree. It determines whether the phenotypic locus is transmitted together with genetic markers of known chromosomal position. Association studies determine whether a certain allele occurs at a frequency higher than that expected by chance in individuals with a particular phenotype. Such an association is thus suggested by a statistically significant difference in the prevalence of alleles with respect to the phenotype. Association studies consisted of two strategies: the candidate gene approach and the genome-wide approach. The candidate gene approach involves the direct examination of whether an individual gene or genes might contribute to the trait of interest. This strategy has been widely applied to analysis of the possible association between genetic variants and disease outcome, with genes selected on the basis of a priori hypotheses regarding their potential etiologic role. It is characterized as a hypothesis-testing approach because of the biological observation supporting the proposed candidate gene. The candidate gene approach is not able, however, to identify disease-associated polymorphisms in unknown genes. In the genome-wide scan, single nucleotide polymorphisms (SNPs) or copy number variations (CNVs) distributed throughout the entire genome are used to identify genomic regions that harbor genes that influence the trait of interest with a detectable effect size. This is a hypothesis-generating approach, allowing the detection of previously unknown potential trait loci.

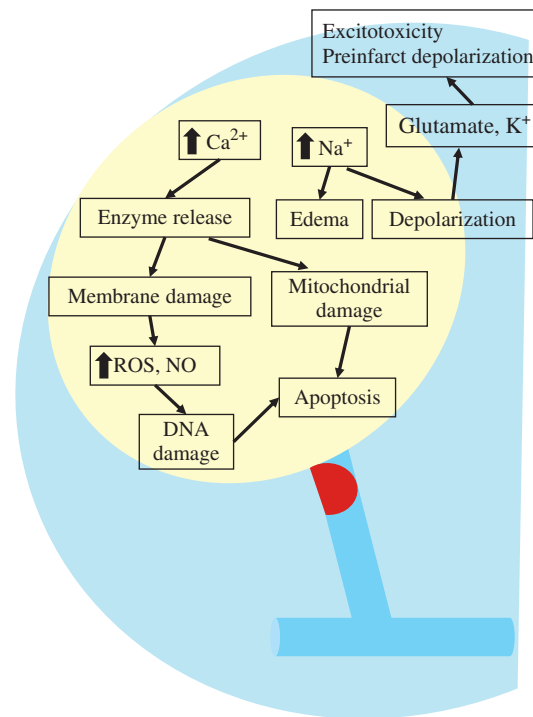
The recent development of high-density genotyping arrays has improved the resolution of unbiased genome-wide scans for common variants associated with multifactorial diseases. Currently, the genome-wide association study (GWAS) makes use of high-throughput genotyping technologies that include up to 4.3 million markers for SNPs and CNVs to examine their relation to clinical conditions or measurable traits. Until February 25, 2012 a Catalog of Published Genome-Wide Association Studies (National Human Genome Research Institute, NIH; <http://www.genome.gov/gwastudies/>) includes 1183 publications and 5910 SNPs associated with various diseases or traits, many in genes not previously suspected of having a role in the condition studied, and some in genomic regions containing no known genes. GWASs represent a substantial advance in the search for genetic variants that confer susceptibility to multifactorial polygenic diseases. GWASs, however, had disadvantages that previously available marker sets were designed to identify common alleles and were not well suited to study the effects of rare variants within a gene of interest.

### 12.4.2 Molecular pathophysiology of ischemic stroke

Stroke is divided into two major varieties, ischemic and hemorrhagic strokes, with most (~87%) cases being ischemic. Ischemic stroke, which includes atherothrombotic cerebral infarction, cardioembolic stroke, and lacunar infarction (Fig. 12.3), is characterized by a sudden decrease in blood flow to one or more central nervous system



**FIGURE 12.3** Etiology of ischemic stroke. Several risk factors have been shown for atherothrombotic cerebral infarction, cardioembolic stroke, or lacunar infarction, including hypertension, diabetes mellitus, and smoking. In addition to these conventional risk factors, genetic factors and interactions between multiple genes and environmental factors are important in the development of ischemic stroke.

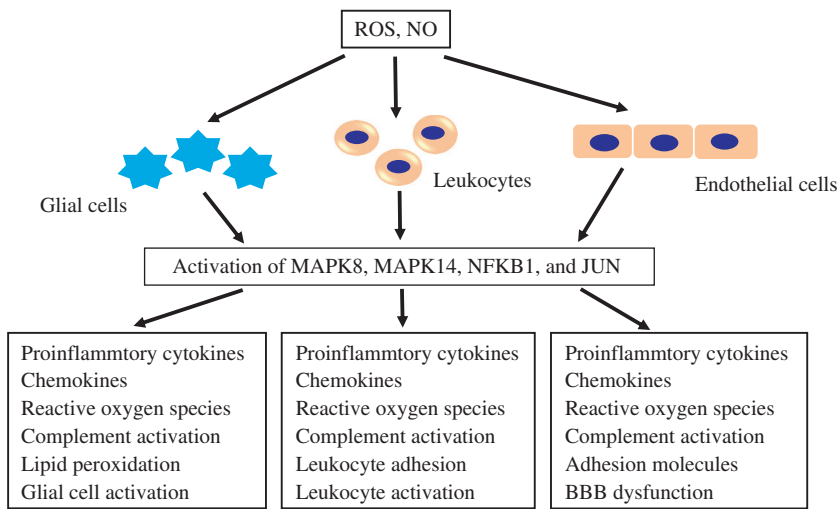


**FIGURE 12.4** Proposed mechanisms of brain injury after ischemic stroke [41]. Ischemia-induced energy failure leads to the depolarization of neurons. Activation of specific glutamate receptors dramatically increases intracellular  $\text{Ca}^{2+}$  and  $\text{Na}^{+}$ , and  $\text{K}^{+}$  is released into the extracellular space. Brain edema results from water shifts to the intracellular space. Increased levels of intracellular  $\text{Ca}^{2+}$  activate proteases, lipases, and endonucleases. ROS are generated and damage membranes, mitochondria, and DNA, resulting in triggering cell death. ROS, Reactive oxygen species.

territories [38] and is a heterogeneous disease caused by different pathogenic mechanisms that include both environmental and genetic factors.

The pathophysiology of ischemic stroke is complex and involves numerous processes, including energy failure, loss of cell ion homeostasis, acidosis, increased intracellular calcium levels, excitotoxicity, free radical-mediated toxicity, generation of arachidonic acid products, cytokine-mediated cytotoxicity, complement activation, disruption of the blood–brain barrier, activation of glial cells, and infiltration of leukocytes [41] (Figs. 12.4 and 12.5). Within a few minutes of a cerebral ischemia, the core of brain tissue, exposed to the most dramatic reduction of blood flow, is mortally injured and, subsequently, undergoes necrotic cell death. This necrotic core is surrounded by a zone of less severely affected tissue which is rendered functionally silent by reduced





**FIGURE 12.5** Proposed molecular pathophysiology of brain injury after ischemic stroke [41]. ROS induces the production of inflammatory mediators, resulting in activation of MAPK8, MAPK14, NFKB1, and JUN in glial cells, endothelial cells, and infiltrating leukocytes. This enhances secretion of proinflammatory cytokines and chemokines and leads to the invasion of leukocytes through upregulation of endothelial adhesion molecules. BBB, Blood–brain barrier; JUN, Jun proto-oncogene; MAPK14, mitogen-activated protein kinase 14; MAPK8, mitogen-activated protein kinase 8; ROS, reactive oxygen species.

blood flow but remains metabolically active [42,43]. Necrosis is morphologically characterized by initial cellular and organelle swelling, subsequent disruption of nuclear, organelle, and plasma membranes, disintegration of nuclear structure, and cytoplasmic organelles with extrusion of cell contents into the extracellular space [42,43]. The region bordering the infarction core, known as the ischemic penumbra, comprises as much as half of the total lesion volume during the initial stages of ischemia and represents the region in which there is opportunity for salvage by poststroke therapy [44]. Less severe ischemia, as occurs in the penumbra region of a focal infarction, evolves more slowly and depends on the activation of specific genes and may ultimately result in apoptosis [45–47]. Recent studies revealed that many neurons in the ischemic penumbra or periinfarction zone may undergo apoptosis after several hours or days, and thus they are potentially recoverable for some time after the onset of stroke. In contrast to necrosis, apoptosis appears to be an orderly process of energy-dependent programmed cell death to dispose of redundant cells. Cells undergoing apoptosis are dismantled in an organized way that minimizes damage and disruption to neighboring cells [43]. There are two general pathways for activation of apoptosis: intrinsic and extrinsic pathways [41].

In rodent stroke models, neurons in the ischemic penumbra show morphological and molecular changes consistent with apoptosis, including caspase activation, expression of proapoptotic genes, and release of cytochrome *c* [47]. Signaling pathways involving hydrolysis of membrane phospholipids are implicated in neuronal apoptosis in stroke [48]. Cleavage of membrane sphingomyelin by acidic sphingomyelinase generates the lipid mediator ceramide. Focal cerebral ischemia in mice induces large increases in acidic sphingomyelinase activity and ceramide levels and the production of inflammatory cytokines [49]. In mice lacking acidic sphingomyelinase the production of cytokines is suppressed, brain damage is decreased, and symptoms are improved [49]. Mice lacking phospholipase A2 also show decreased brain damage after focal cerebral ischemia, suggesting an important function of lipid mediators generated by this enzyme in ischemic neuronal injury [50].

### 12.4.3 Molecular genetics of ischemic stroke

Studies with twins, siblings, and families have provided substantial evidence for stroke heritability [3]. A mechanistic approach to the study of ischemic stroke, as advocated in the Trial of ORG 10172 in Acute Stroke Treatment study [51], is the best suited to genetic research [52]. This approach classifies ischemic stroke into five subtypes: (1) large-artery atherosclerosis; (2) small-vessel occlusion; (3) cardiogenic embolism; (4) stroke of other determined etiology; and (5) stroke of undetermined etiology. It was applied in a family history study of 1000 individuals with ischemic stroke and 800 controls [53]. This study found that a family history of vascular disease was a risk factor for both small-vessel occlusion and large-vessel atherosclerosis but not for cardioembolic stroke or stroke of undetermined etiology. These findings suggest that genetic research may be most fruitful when focused on the former two subtypes of ischemic stroke [52].

Accurate phenotyping and performance of separate analyses according to stroke subtypes are thus essential. Focusing on particular stroke subtypes will likely make a study more efficient and markedly reduce the necessary sample sizes [53]. Another way of increasing the statistical power of a study may be to focus on early-onset



cases, as is in any genetic predisposition. Family history studies, such as prospective twin studies [54], suggest that the genetic component of stroke is stronger in such individuals.

The main cause of ischemic stroke is atherothrombosis, with the principal and treatable risk factors including hypertension, diabetes mellitus, and dyslipidemia [55]. In addition to these conventional risk factors, genetic variants are important in the pathogenesis of ischemic stroke [7,56]. Prediction of the risk for ischemic stroke beyond the usual clinical risk factors on the basis of genetic variants would be useful for deciding how aggressively to target the risk factors that are currently amenable to treatment. Furthermore, it might prompt earlier carotid imaging of patients at risk in order to detect asymptomatic carotid stenosis [56].

A whole-genome linkage analysis of families or sibling pairs showed that chromosomal region 5q12 was linked to ischemic stroke [57]. A large number of candidate gene association studies of unrelated individuals have identified many genes that are related to the prevalence of ischemic stroke (Table 12.1). Candidate gene association studies, however, have substantial limitations for detecting the genetic basis of stroke because this approach relies on selection of the genes for association studies based on either a biological hypothesis or the location of a particular gene in implicated linkage regions. In addition, most candidate gene association studies for stroke have generated inconsistent or inconclusive results. Recent GWASs identified several loci and genes that confer susceptibility to ischemic stroke. The published results of GWASs for ischemic stroke are summarized in Table 12.2. In the following section, candidate genes for ischemic stroke of particular interest [*PDE4E*, arachidonate 5-lipoxygenase-activating protein gene (*ALOX5AP*), *CDKN2B* antisense RNA 1 gene (*CDKN2B-AS1*), *Ninjurin 2* gene (*NINJ2*), paired-like homeodomain 2 gene (*PITX2*), and zinc finger homeobox 3 gene (*ZFXH3*)] are reviewed.

**TABLE 12.1** Genes shown to be related to ischemic stroke by linkage analyses or candidate gene association studies.

Chromosomal locus	Gene name	Gene symbol	References
1p36.3	Methylenetetrahydrofolate reductase	<i>MTHFR</i>	[58]
1p36.2	Natriuretic peptide A	<i>NPPA</i>	[59]
1q21-q23	C-reactive protein, pentraxin-related	<i>CRP</i>	[60]
1q23-q25	Selectin P	<i>SELP</i>	[61]
1q25.2-q25.3	Prostaglandin-endoperoxide synthase 2	<i>PTGS2</i>	[62]
2q14	Interleukin 1, beta	<i>IL1B</i>	[63]
2q14.2	Interleukin 1 receptor antagonist	<i>IL1RN</i>	[64]
3pter-p21	Chemokine (C-X3-C motif) receptor 1	<i>CX3CR1</i>	[65]
3p25	Peroxisome proliferator-activated receptor gamma	<i>PPARG</i>	[66]
4p16.3	Adducin 1	<i>ADD1</i>	[67]
4q28	Fibrinogen beta chain	<i>FGB</i>	[68]
4q28-q31	Fatty acid binding protein 2	<i>FABP2</i>	[69]
5q12	Phosphodiesterase 4D, cAMP-specific	<i>PDE4D</i>	[70]
5q23-q31	Integrin, alpha 2	<i>ITGA2</i>	[71]
5q31.1	Interleukin 4	<i>IL4</i>	[61]
5q32-q33.1	Glutathione peroxidase 3	<i>GPX3</i>	[72]
5q33-qter	Coagulation factor XII	<i>F12</i>	[73]
6p25-p24	Coagulation factor XIII, A1 polypeptide	<i>F13A1</i>	[74]
6p21.3	Lymphotoxin alpha	<i>LTA</i>	[75]
6q22	c-Ros oncogene 1, receptor tyrosine kinase	<i>ROS1</i>	[76]
6q25.1	Estrogen receptor 1	<i>ESR1</i>	[77]

(Continued)

TABLE 12.1 (Continued)

Chromosomal locus	Gene name	Gene symbol	References
6q27	Lipoprotein, Lp(a)	<i>LPA</i>	[78]
7p21	Interleukin 6	<i>IL6</i>	[79,80]
7q21.3	Paraoxonase 1	<i>PON1</i>	[81]
7q21.3-q22	Serpin peptidase inhibitor, clade E, member 1	<i>SERPINE1</i>	[82]
7q36	Nitric oxide synthase 3	<i>NOS3</i>	[83]
8p22	Lipoprotein lipase	<i>LPL</i>	[84]
8p21-p12	Epoxide hydrolase 2, cytoplasmic	<i>EPHX2</i>	[85]
8p12	Plasminogen activator, tissue	<i>PLAT</i>	[86]
9p21.3	CDKN2B antisense RNA 1	<i>CDKN2B-AS1</i>	[87]
9q31.1	ATP-binding cassette, subfamily A, member 1	<i>ABCA1</i>	[76]
11p11	Coagulation factor II	<i>F2</i>	[88]
11q23	Apolipoprotein A-V	<i>APOA5</i>	[89]
12p13	Guanine nucleotide binding protein, beta polypeptide 3	<i>GNB3</i>	[67]
12p13	Sodium channel, nonvoltage-gated 1, alpha	<i>SCNN1A</i>	[90]
13q12	Arachidonate 5-lipoxygenase-activating protein	<i>ALOX5AP</i>	[91]
14q11.2	Cathepsin G	<i>CTSG</i>	[92]
14q22	Prostaglandin E receptor 2, 53 kDa	<i>PTGER2</i>	[93]
16p11.2	Vitamin K epoxide reductase complex, subunit 1	<i>VKORC1</i>	[94]
16q24	Cytochrome b-245, alpha polypeptide	<i>CYBA</i>	[95]
17pter-p12	Glycoprotein Ib, alpha polypeptide	<i>GP1BA</i>	[96]
17q21.32	Integrin, beta 3	<i>ITGB3</i>	[97]
17q23	Angiotensin I converting enzyme	<i>ACE</i>	[98]
19p13.3	Thromboxane A2 receptor	<i>TBXA2R</i>	[99]
19p13.3-p13.2	Intercellular adhesion molecule 1	<i>ICAM1</i>	[79]
19p13.2	Low density lipoprotein receptor	<i>LDLR</i>	[100]
19q13.1	Transforming growth factor, beta 1	<i>TGFB1</i>	[101]
19q13.2	Apolipoprotein E	<i>APOE</i>	[68]
Xq28	Interleukin-1 receptor-associated kinase 1	<i>IRAK1</i>	[76]

#### 12.4.3.1 Phosphodiesterase 4D, cAMP-specific gene

A genome-wide linkage study indicated that a gene on 5q12 may contribute to the risk of stroke [57]. A case-control study was performed to determine which linkage disequilibrium block within the linkage peak showed the strongest association with ischemic stroke. Markers in the alternative promoter region corresponding to one of the eight isoforms of phosphodiesterase 4D (PDE4D) showed the strongest association [70]. The subtypes of ischemic stroke with the highest risk ratios were large-vessel occlusive disease and cardioembolic stroke; there was no association with small-vessel occlusive disease. The frequency of the most significant haplotype in each of these two patient subgroups was ~30%, and the relative risk was 1.98 for the two subgroups combined. A mutually exclusive haplotype that conferred protection was present in 28% of control individuals with a relative risk of 0.68. *PDE4D* variants, thus, conferred substantial risk for two forms of ischemic stroke that are related to atherosclerosis [38]. Neither the risk nor protective haplotypes were associated with underlying missense or nonsense mutations, but they did correlate with the expression of *PDE4D* [70].

**TABLE 12.2** Chromosomal loci and genes shown to be related to ischemic stroke by genome-wide association studies.

Chromosomal locus	dbSNP	Nucleotide substitution	Gene (nearby gene)	Phenotype	References
4q25	rs2200733	C→T	<i>PITX2</i>	Cardioembolic stroke	[102]
4q25	rs1906599	T→C	<i>PITX2</i>	Cardioembolic stroke	[103]
7p21.1	rs11984041	C→T	<i>HDAC9</i>	Cerebral infarction	[103]
11q12	rs9943582	C→T	<i>APLN</i>	Cerebral infarction	[104]
12p13	rs12425791	G→A	<i>NINJ2</i>	Cerebral infarction	[105]
12p13	rs11833579	G→A	<i>NINJ2</i>	Cerebral infarction	[105]
14q23.1	rs2230500	G→A (Val374Ile)	<i>PRKCH</i>	Lacunar infarction	[106]
16q22	rs7193343	T→C	<i>ZFHX3</i>	Cardioembolic stroke	[107]
16q22.3	rs12932445	T→C	<i>ZFHX3</i>	Cardioembolic stroke	[103]
20p12.1	rs2208454	G→T	<i>MACROD2</i>	Cerebral infarction	[108]
22q13.3	rs6007897	A→G (Thr2268Ala)	<i>CELSR1</i>	Cerebral infarction	[109]
22q13.3	rs4044210	A→G (Ile2107Val)	<i>CELSR1</i>	Cerebral infarction	[109]

PDE4D degrades the second messenger cAMP [110], which is a key signaling molecule in cell types that are important in the pathogenesis of atherosclerosis [38]. A decrease in cAMP levels in vascular smooth muscle cells in vitro promoted the proliferation and migration of these cells, processes that are characteristic of atherosclerosis [110–113]. Inhibitors of PDE4 were found to block smooth muscle proliferation induced in the rat carotid artery by balloon injury [114,115]. *PDE4D* is also expressed in activated macrophages and may therefore play a role in inflammation within atherosclerotic plaques, possibly contributing to atherogenesis or plaque instability, or both [116–118]. Increased activity of one or more isoforms of PDE4D resulting from dysregulation of transcript splicing or translation may thus increase the risk for ischemic stroke, with the decreased risk conferred by the identified protective haplotype possibly being due to a reduced activity of PDE4D [38].

Studies that have attempted to replicate this association of *PDE4D* with stroke have yielded diverse results [119]. In a UK population, no overall association was found with ischemic stroke, but possible associations were identified with cardioembolic stroke and large-artery stroke [120]. A US study reported an association of *PDE4D* with ischemic stroke, especially with large-artery stroke [121]. In contrast, no association was found in a German stroke cohort [122] or a Swedish stroke cohort of individuals aged <75 years [123]. A linkage study with a second Swedish population confirmed linkage of ischemic stroke to 5q12 [123], but no linkage was detected in an American population [121]. No association of *PDE4D* was found with carotid intimal-medial thickness [120], suggesting that the gene does not exert its effects by accelerating early atherosclerosis [119]. A metaanalysis comprising 16 studies in 5216 cases and 6615 controls failed to detect the relation of *PDE4E* variants with ischemic stroke [124]. A multilocus Bayesian metaanalysis including 14 data sets from populations of European descent and genotypes of 33 SNPs in 12,929 subjects (5994 cases and 6935 controls) confirmed no association despite the increase in statistical power [125]. Although the regulation of intracellular cAMP concentration by PDE4D in vascular smooth muscle cells or even in macrophages may be a key determinant of stroke risk, the role of *PDE4D* variants in the pathogenesis of ischemic stroke remains unclear.

#### 12.4.3.2 Arachidonate 5-lipoxygenase-activating protein gene

A linkage and association study in Iceland demonstrated that *ALOX5AP* confers risk for both myocardial infarction and ischemic stroke [91]. The locus associated with myocardial infarction was initially mapped to chromosome 13q12 through a genome-wide linkage scan conducted on 296 families with this condition [91]. An independent linkage study of Icelandic stroke patients without myocardial infarction identified the same locus [91]. The haplotype defined by microsatellite markers that showed the strongest association with myocardial infarction covered a region containing *ALOX5AP* [38]. A haplotype that spans *ALOX5AP* and is defined by four SNPs (HapA) was subsequently shown to be associated with myocardial infarction, with a relative risk of 1.8. The same haplotype was then found to confer risk for stroke in the Icelandic population with a relative risk

of 1.7 [38]. HapA is relatively common and is carried by 27% of Icelandic patients with stroke. Another haplotype within *ALOX5AP* (HapB) showed a significant association with myocardial infarction in British cohorts, with a relative risk of 2.0 [91]. The association of both haplotypes with stroke in Scottish individuals was subsequently demonstrated [126].

*ALOX5AP* participates in the initial steps of leukotriene synthesis. Arachidonic acid is thus converted to leukotriene A4 by the action of arachidonate 5-lipoxygenase and its activating protein, *ALOX5AP*. Inflammatory lipid mediators, including leukotrienes B4, C4, D4, and E4 [127], are then produced from leukotriene A4 by the action of leukotriene A4 hydrolase and leukotriene C4 synthase. The amount of leukotriene B4 synthesized by ionomycin-stimulated neutrophils from individuals with myocardial infarction was greater than that produced by those from control individuals [91], supporting the notion that the increased activity of the leukotriene pathway plays a role in the pathogenesis of myocardial infarction [38]. Moreover, the observed difference in the release of leukotriene B4 was largely accounted for by carriers of HapA, whose cells produced more leukotriene B4 than did those from noncarriers. Although leukotriene B4 production was not measured in cells from patients with stroke, a similar increase would be expected, given that the HapA variant of *ALOX5AP* shows similar associations with ischemic stroke and myocardial infarction. Elevated levels of leukotriene B4 might contribute to atherogenesis or plaque instability by promoting inflammation at atherosclerotic plaques [38].

A role for upregulation of the leukotriene pathway in atherosclerosis is further supported by the observation that expression of enzymes of the arachidonate 5-lipoxygenase pathway is increased in human atheromas, with the number of arachidonate 5-lipoxygenase–positive cells (macrophages, dendritic cells, mast cells, and neutrophils) being markedly increased in advanced lesions [128]. Furthermore, the arachidonate 5-lipoxygenase gene (*ALOX5*) has been implicated in the development of atherosclerosis in mice by the finding that the loss of only one *ALOX5* allele confers protection against atherosclerosis in animals deficient in the low density lipoprotein receptor [129]. An increased activity of the leukotriene biosynthetic pathway associated with specific *ALOX5AP* variants might thus promote the processes of atherogenesis and subsequent plaque instability, increasing the chance of ischemic stroke on the background of atherosclerosis [38].

Several groups have attempted to replicate the association of ischemic stroke with *ALOX5AP* variants [119]. Although the deCODE group replicated the association in a Scottish stroke population [126], a case-control study in Germany reported a weak association with an *ALOX5AP* polymorphism [122], and no association in a case-control study and no linkage to this chromosomal region in a sibling-pair study were found in American populations [120]. A metaanalysis of 5194 stroke cases and 4566 controls showed significant heterogeneity among studies and failed to detect overall significant association [130]. Although the leukotriene biosynthetic pathway may play an important role in leukocyte chemotaxis and inflammatory responses that are key processes in atherosclerosis, the role of *ALOX5AP* polymorphisms in the pathogenesis of stroke remain elucidated.

#### 12.4.3.3 *CDKN2B antisense RNA 1 gene*

Chromosome 9p21.3 locus was first discovered by GWASs to be a risk factor for coronary heart disease or myocardial infarction [131–134]. This locus was then shown to be associated with ischemic stroke [135,136]. A metaanalysis of eight studies (9632 cases and 30,716 controls) confirmed the association of rs10757278 SNP at 9p21.3 with ischemic stroke. Similar analysis of two studies (5255 cases and 22,640 controls) also demonstrated the relation of rs1537378 at 9p21.3 to ischemic stroke [87]. The locus was also associated with intracranial and aortic aneurysms [137]. Recently, 9p21.3 has also been associated with platelet reactivity [138]. The increased platelet reactivity may explain the association with myocardial infarction and ischemic stroke; however, it is not clear how increased platelet reactivity might relate to the formation of aneurysm. The 9p21.3 locus includes the *CDKN2B-AS1*, which alters the expression of several genes related to cellular proliferation [139]. Recent study also indicates that 9p21 polymorphisms influence inflammatory signaling [140] and vascular cell proliferation [141]. These studies suggest that *CDKN2B-AS1* at 9p21.3 may be a susceptibility locus for myocardial infarction and ischemic stroke, although the underlying molecular mechanisms have not been determined definitively. An improved explanation for the pleiotropic effects of the 9p21.3 locus on multiple vascular beds should emerge in the near future [142].

#### 12.4.3.4 *Ninjurin 2 gene*

A prospective GWAS showed an association of two SNPs on chromosome 12p13 with ischemic stroke [105]. The primary study population, consisting of 19,602 subjects including 1544 incident strokes from four prospective white cohorts, identified two SNPs (rs11833579 and rs12425791) in the region of 12p13. Results of the initial GWAS suggested that a minor allele of each SNP increased the hazard ratio for total stroke by ~1.3 and for

ischemic stroke by  $\sim 1.4$ . The corresponding population-attributable risks were 11%–13% for total stroke and 14%–17% for ischemic stroke. The association between the two SNPs and ischemic stroke was further tested in two independent replication samples, an African-American community-based cohort and a Dutch case-control sample, and rs12425791 was significantly associated with ischemic stroke, especially atherothrombotic stroke in both samples. This SNP was located close to *NINJ2*, which encodes an adhesion molecule expressed in glia that shows increased expression after nerve injury [105]. A metaanalysis combining data for 8637 cases and 8733 controls of European ancestry and a population-based genome-wide cohort study of 278 ischemic stroke among 22,054 participants did not replicate the relation of either SNP with ischemic stroke [143]. In addition, Italian [144] and Swedish [145] case-control studies also failed to show any association between the two genetic variants on 12p13 and ischemic stroke [142]. The relation of polymorphisms at 12p13 locus to ischemic stroke and the underlying molecular mechanisms thus remain unclear.

#### 12.4.3.5 Paired-like homeodomain 2 gene and zinc finger homeobox 3 gene

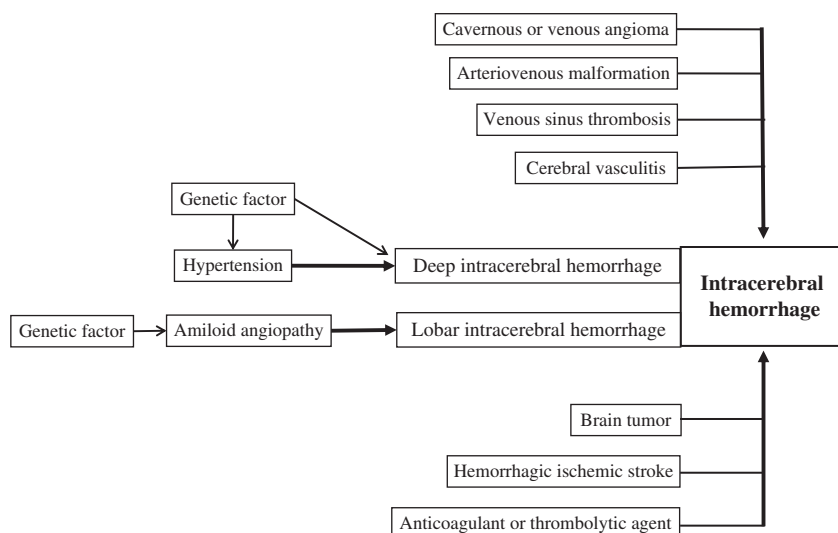
Atrial fibrillation is an important risk factor for cardioembolic stroke, and there are overlapping genetic risk factors for these two conditions. A multistage study in 1661 cases with ischemic stroke and 10,815 controls in the discovery phase showed that two SNPs (rs2200733 and rs10033464) at 4q25 near *PITX2* were significantly associated with cardioembolic stroke [102]. These SNPs had already been associated with atrial fibrillation. Another association study involving 4199 cases of ischemic stroke and 3750 controls showed an association between the 4q25 locus and cardioembolic stroke [146].

Overlapping genetic risk for atrial fibrillation and ischemic stroke was further observed for a sequence variant in the zinc finger homeobox 3 gene (*ZFHX3*) on 16q22 [107]. The GWAS of samples from Iceland, Norway, and the United States observed that rs7193343 was significantly associated with atrial fibrillation. This variant was also associated with cardioembolic stroke. *ZFHX3* encodes a transcription factor with multiple homeodomains and zinc finger motifs and regulates myogenic and neuronal differentiation [26].

#### 12.4.4 Molecular pathophysiology of intracerebral hemorrhage

Intracerebral hemorrhage is responsible for  $\sim 10\%$  of all strokes, including a large proportion of fatal or severe cases. Advancing age and hypertension are the most important risk factors for intracerebral hemorrhage. Intracerebral hemorrhage is usually attributed to hypertensive small-vessel disease, with the most common sites of hemorrhage being the basal ganglia, cerebellum, and pons (Fig. 12.6).

Intraparenchymal bleeding results from the rupture of the small penetrating arteries that originate from basilar arteries or the anterior, middle, or posterior cerebral arteries. Degenerative changes in the vessel wall induced by chronic hypertension reduce compliance and increase the likelihood of spontaneous rupture. Electron-microscopical studies suggest that most bleeding occurs at or near the bifurcation of affected arteries, where prominent degeneration of the media and smooth muscles can be observed [147,148].

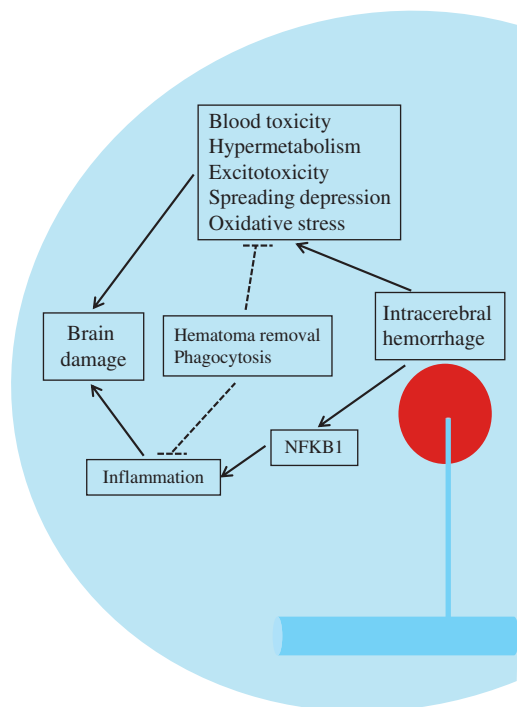


**FIGURE 12.6** Etiology of intracerebral hemorrhage. Intracerebral hemorrhage includes hypertensive deep cerebral hemorrhage and lobar cerebral hemorrhage caused by amyloid angiopathy. Intracerebral hemorrhage is also occurred by various conditions, including cavernous or venous angioma, arteriovenous malformation, venous sinus thrombosis, cerebral vasculitis, brain tumor, hemorrhagic ischemic stroke, and anticoagulant or thrombolytic agent. Although hypertension is the most important risk factor for intracerebral hemorrhagic, genetic factors are also involved in the development of this condition.



During intracerebral hemorrhage, rapid accumulation of blood within brain parenchyma leads to disruption of normal anatomy and an increase in local pressure. Depending on the expansion of hematoma, the primary damage occurs within minutes to hours from the onset of bleeding and is primarily the result of mechanical damage associated with the mass effect [149]. Secondary damage is attributable to the presence of intraparenchymal blood and may depend on the initial hematoma volume, age, or ventricular volume [149]. Secondary damage may occur through many parallel pathological pathways, including cytotoxicity of blood [150,151], hypermetabolism [152], excitotoxicity [153], spreading depression [154], and oxidative stress and inflammation [151,155–161] (Fig. 12.7). These pathological events lead to irreversible disruption of the components of the neurovascular unit, constituting gray and white matters, and is followed by disruption of blood–brain barrier and deadly brain edema with massive brain cell death [150,151,153,155,163–165]. Although inflammatory mediators generated locally in response to brain injury augment damage caused by intracerebral hemorrhage (secondary injury), the involvement of inflammatory cells, such as microglia and macrophages, is vital for the removal or cleanup of cellular debris from hematoma, the source of ongoing inflammation [166]. The timely removal of damaged tissue is essential for reducing the length of deleterious pathological process, thereby allowing faster and more efficient recovery [162].

The inflammatory signaling in a hemorrhage-affected brain involves a transcription factor, nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB1) [161,167]. The target genes of NFKB1 include those for various adhesion molecules including intercellular adhesion molecule 1; proinflammatory cytokines including interleukin 1, beta and tumor necrosis factor; chemokines; matrix metalloproteinases including matrix metalloproteinase 9; immune receptors; acute phase proteins; cell surface receptors; and inflammatory enzymes including nitric oxide synthase 2, inducible, prostaglandin-endoperoxide synthase 2, and phospholipase A2 [162]. Reactive oxygen species act as important signaling molecules in the activation of NFKB1. This property of NFKB1 may, in part, explain how oxidative stress enhances inflammation after intracerebral hemorrhage [162]. Experimental studies demonstrate that NFKB1 is activated in a hemorrhage-affected hemisphere as early as 15 minutes after the onset of intracerebral hemorrhage, reaches maximum between 1 and 3 days, and remains elevated for weeks



**FIGURE 12.7** Proposed mechanisms of secondary brain injury after intracerebral hemorrhage [162]. Intracerebral hemorrhage activates NFKB1, which then induces inflammation that leads to secondary brain damage. Intracerebral hemorrhage also induces blood toxicity, hypermetabolism, excitotoxicity, spreading depression, and oxidative stress, which result in brain damage. Hematoma removal and phagocytosis suppress these adverse events, leading to the reduction of brain damage. *NFKB1*, Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1.



[167]. Interleukin 1, beta, tumor necrosis factor, and matrix metalloproteinase 9, which are upregulated by NFKB1, are shown to be involved in hemorrhage-mediated brain injury [157,168–171].

### 12.4.5 Molecular genetics of intracerebral hemorrhage

Familial aggregation of cases of intracerebral hemorrhage was demonstrated in a prospective study in North Carolina in the United States, which found that 10% of affected individuals had a family history of intracerebral hemorrhage [172]. No significant clinical demographic differences separated affected individuals with or without a family history of intracerebral hemorrhage. Genetic factors may influence not only the development of intracerebral hemorrhage but also the prevalence of certain risk factors for this condition, such as hypertension. Furthermore, such genetic factors may interact with environmental factors such as diet and cigarette smoking.

In some individuals with intracerebral hemorrhage the hemorrhage is lobar in location, such as in the frontal, parietal, temporal, or occipital cortex, and such patients often do not have hypertension [173]. This category of hemorrhage, referred to as lobar intracerebral hemorrhage, may represent a distinct pathogenetic subgroup [174] (Fig. 12.6). The occurrence of lobar intracerebral hemorrhage was shown to be associated with the  $\epsilon 2$  and  $\epsilon 4$  alleles of the apolipoprotein E gene (*APOE*) [175,176]. These relations, particularly that with  $\epsilon 4$ , are presumably due to the association of *APOE* with cerebral amyloid angiopathy [177].

Various candidate gene association studies of unrelated individuals have identified genes that are related to intracerebral hemorrhage (Table 12.3). The etiology of intracerebral hemorrhage is complex and the genetic determinants of this condition are still largely unknown. In the following sections, candidate genes for intracerebral hemorrhage that are of particular interest (*APOE* and *COL4A1*) are reviewed.

#### 12.4.5.1 Apolipoprotein E gene

Cerebral amyloid angiopathy is a frequent cause of lobar intracerebral hemorrhage [174]. The main pathological feature of this condition is the infiltration of cortical vessels by amyloid beta, a homogeneous eosinophilic substance found in the brain of elderly individuals and an important component of the senile plaques in patients with Alzheimer's disease. The incidence of lobar intracerebral hemorrhage due to cerebral amyloid angiopathy increases markedly with age, with most affected individuals being over the age of 60 and many having antecedent memory loss. Patients with cerebral amyloid angiopathy and intracerebral hemorrhage have a lower mortality rate and a greater risk of recurrence than those with other types of intracerebral hemorrhage [174].

The  $\epsilon 4$  and  $\epsilon 2$  alleles of *APOE* were identified as predictors of recurrent lobar intracerebral hemorrhage in patients with cerebral amyloid angiopathy [175]. The risk of recurrence at 2 years was 28% for carriers of the  $\epsilon 2$  or  $\epsilon 4$  alleles, compared with 10% for patients with the  $\epsilon 3$  allele. The presence of the  $\epsilon 2$  or  $\epsilon 4$  alleles is thus considered a potent risk factor for recurrence [174].

**TABLE 12.3** Genes shown to be related to intracerebral hemorrhage by candidate gene association studies.

Chromosomal locus	Gene name	Gene symbol	References
6p25-p24	Coagulation factor XIII, A1 polypeptide	<i>F13A1</i>	[178]
6q27	Lipoprotein, Lp(a)	<i>LPA</i>	[78]
7p21	Interleukin 6	<i>IL6</i>	[80]
7q11.23	Lim domain kinase 1	<i>LIMK1</i>	[76]
9q34.1	Endoglin	<i>ENG</i>	[179]
13q34	Collagen, type IV, alpha 1	<i>COL4A1</i>	[180]
14q32.1	Serpin peptidase inhibitor, clade A, member 3	<i>SERPINA3</i>	[181]
16p11.2	Vitamin K epoxide reductase complex, subunit 1	<i>VKORC1</i>	[94]
17q23	Angiotensin I converting enzyme	<i>ACE</i>	[182]
17q23-qter	Apolipoprotein H	<i>APOH</i>	[183]
19q13.2	Apolipoprotein E	<i>APOE</i>	[175]

Other studies have shown that the frequency of the  $\epsilon 4$  allele of *APOE* is increased in patients with cerebral amyloid angiopathy, whereas the  $\epsilon 2$  allele may be associated with an increased risk of intracerebral hemorrhage in individuals with this condition [175,184–186]. These observations may be partially explained by the association of the  $\epsilon 4$  allele with Alzheimer's disease, which is present in up to 50% of patients with cerebral amyloid angiopathy. In addition the age of onset of intracerebral hemorrhage in individuals with cerebral amyloid angiopathy was found to be earlier in carriers of the  $\epsilon 4$  allele of *APOE* than in carriers of other alleles of this gene [177].

#### 12.4.5.2 Collagen, type IV, alpha 1 gene

Mutations in *COL4A1* have been detected in families with cerebral small-vessel disease [180,187]. *COL4A1* was initially identified as the causative gene in a mouse mutant with perinatal cerebral hemorrhage and porencephaly [187]. Mice heterozygous for the responsible mutation develop recurrent hemorrhage in the basal ganglia, a typical site of intracerebral hemorrhage in hypertensive patients. Subsequent analysis of families with porencephaly and cerebral small-vessel disease revealed several mutations of human *COL4A1* in affected individuals [180,187,188].

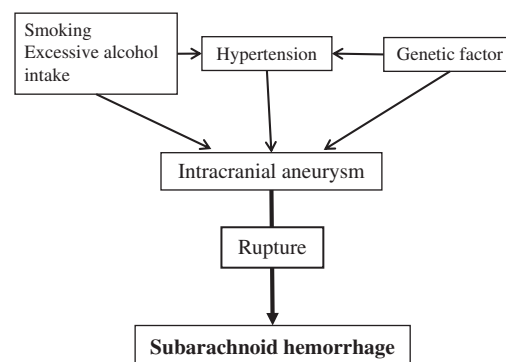
Type IV collagens are an integral component of the vascular basement membrane. *COL4A1* and *COL4A2*, the most abundant type IV collagens, form heterotrimers. Repeated Gly-Pro-X motifs of these collagen molecules are required for the formation of a triple helix during collagen assembly, and most mutations identified in *COL4A1* affect Gly residues within these motifs. It was therefore hypothesized that *COL4A1* mutations interfere with triple-helix formation or heterotrimer secretion. Indeed, analysis of heterozygous embryonic tissue has suggested that the mutations inhibit collagen secretion into the basement membrane. Ultrastructural abnormalities in capillaries of carriers of *COL4A1* mutations are indicative of disordered assembly of the basement membrane [189].

The phenotypic spectrum associated with *COL4A1* mutations is broad and strongly related to small-vessel disease. Key features include leukoencephalopathy, microhemorrhages, and clinically overt hemorrhage. The structural compromise of small blood vessels is reflected by the observations that birth trauma, brain trauma, or oral anticoagulants may trigger intracerebral hemorrhage in carriers of *COL4A1* mutations [180,187]. Genes that encode proteins associated with the vascular basement membrane are thus potential candidates for causative agents of intracerebral hemorrhage and leukoencephalopathy [189].

#### 12.4.6 Molecular pathophysiology of intracranial aneurysm and subarachnoid hemorrhage

Subarachnoid hemorrhage is most commonly caused by the rupture of an aneurysm on an intracranial artery [190]. About 2% of the general population has an intracranial aneurysm [191]. Rupture of an aneurysm is most common between 40 and 60 years of age and prognosis after rupture is poor, with half of affected individuals dying within 1 month and 20% remaining dependent on support for activities of daily life [192–194]. The incidence of aneurysmal subarachnoid hemorrhage in the general population is low (about 8 per 100,000 person years) [195], but the young age at onset and the poor prognosis mean that the loss of productive life years is similar to that associated with ischemic stroke [196].

The pathogenesis of subarachnoid hemorrhage from a ruptured aneurysm has not been fully understood. Hemodynamic factors and structural properties of the arterial wall may contribute to the development of intracranial aneurysms, but the trigger factors remain unknown (Fig. 12.8). Disruption of the extracellular matrix is

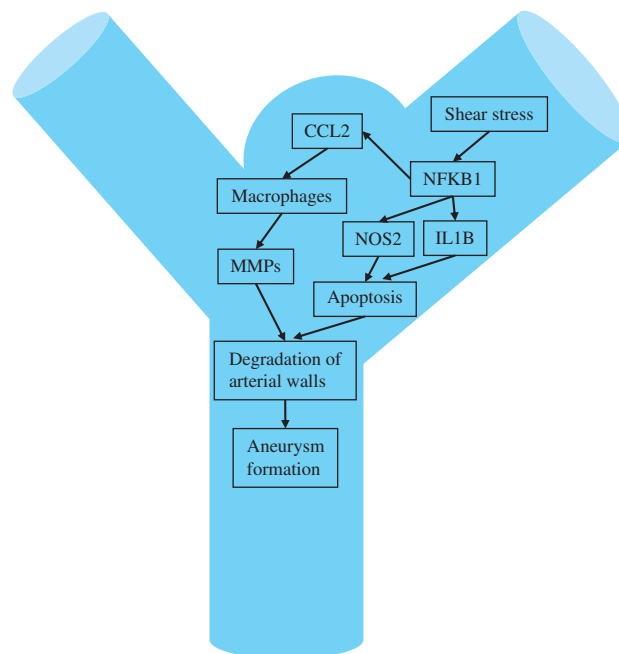


**FIGURE 12.8** Etiology of subarachnoid hemorrhage. Hypertension, smoking, and excessive alcohol intake are risk factors for subarachnoid hemorrhage. In addition to these factors, genetic factors are important in the development of intracranial aneurysm and subarachnoid hemorrhage.

likely to be a factor in the pathophysiology given that intracranial aneurysms are associated with heritable disorders of connective tissue and the extracellular matrix [190,197,198]. Moreover, the amount of structural proteins of the extracellular matrix has been found to be reduced in the intracranial arterial wall of many ruptured intracranial aneurysms as well as in skin biopsies and intracranial and extracranial arteries of patients with aneurysms [199–205].

Recent studies suggested that intracranial aneurysm is a chronic inflammatory disease at the bifurcation site of cerebral arterial walls [206] (Fig. 12.9). Many inflammatory cells were observed in human samples from surgically dissected intracranial aneurysm walls [207]. In the intracranial aneurysms of rodent models, macrophages were the main inflammatory cells, and the number of macrophages gradually increased during aneurysm formation, suggesting the involvement of macrophage-mediated inflammation in the formation of intracranial aneurysm [208]. Chemokine (C-C motif) ligand 2 (CCL2) is the major chemoattractant for macrophages and contributes to recruitment of macrophages to the inflammatory site. The CCL2 mRNA was expressed in human intracranial aneurysm walls, mainly in endothelial cells, at the early stage of aneurysm formation, and CCL2 deficiency resulted in reduced incidence of intracranial aneurysm through suppression of macrophage infiltration to the arterial walls [209]. These observations suggest that macrophages contribute to the formation of intracranial aneurysm and that chronic inflammation is involved in the pathogenesis of intracranial aneurysm [206].

Degeneration of the extracellular matrix is the major pathological feature of intracranial aneurysm. Degeneration of the extracellular matrix may be related to enlargement and rupture of the aneurysm as a result of weakening of arterial walls. Increased collagenolytic activity and reduced collagen content were detected in human intracranial aneurysm walls [210,211]. Matrix metalloproteinases (MMPs), especially MMP9, may have an important role in the formation of intracranial aneurysm [208,212,213]. The expression of MMP9 was increased in macrophages infiltrating aneurysm walls. Inhibition of the MMP9 activity by a selective inhibitor resulted in reduced incidence of intracranial aneurysm, suggesting the important role of MMP9 in the formation of intracranial aneurysm [208]. Apoptosis of medial smooth muscle cells is a possible mechanism of degeneration of the media [214]. Among molecules related to apoptosis cell signaling, nitric oxide synthase 2, nitric oxide synthase 2 (NOS2) and interleukin 1, beta (IL1B) may contribute to the formation of intracranial aneurysm. NOS2 and IL1B were expressed in the media of aneurysm walls. Deficiency of NOS2 or IL1B resulted in reduced aneurysm



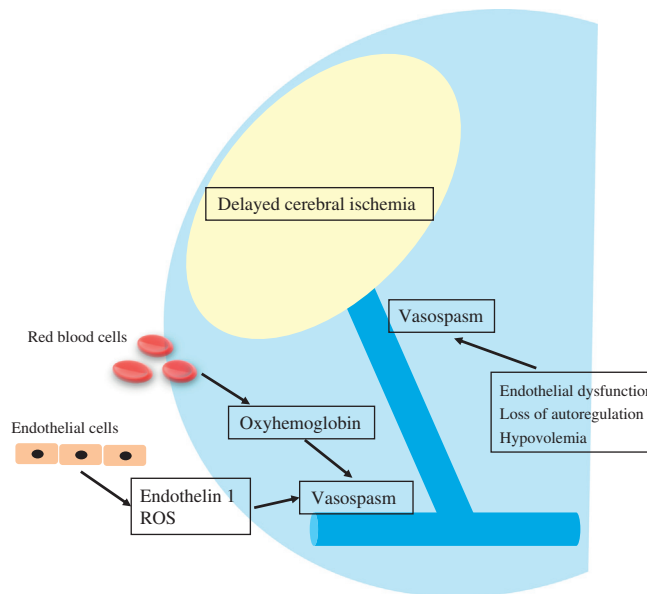
**FIGURE 12.9** Proposed mechanisms of the formation of intracranial aneurysm [206]. Shear stress activates NFKB1 in endothelial cells. Activation of NFKB1 increases the expression of NOS2 and IL1B, resulting in acceleration of apoptosis of arterial wall cells. Activation of NFKB1 also increases release of CCL2 that recruits macrophages to the arterial walls. Macrophages produce MMP, especially MMP9, resulting in degradation of arterial walls. Weakening of arterial walls by chronic inflammation may thus be a crucial mechanism of the formation of intracranial aneurysm. CCL2, Chemokine (C-C motif) ligand 2; IL1B, interleukin 1, beta; MMPs, matrix metalloproteinases; NFKB1, nuclear factor of kappa light polypeptide gene enhancer in B-cells 1; NOS2, nitric oxide synthase 2, inducible.

formation by suppression of apoptosis in medial smooth muscle cells of the arterial walls, suggesting the involvement of apoptotic cell death in the pathogenesis of intracranial aneurysm [215–218].

Various proinflammatory genes, including MMPs, NOS2, and CCL2, have specific NFκB1 binding sites, suggesting the role of NFκB1 in the pathogenesis of aneurysm formation through the transcriptional regulation of proinflammatory genes [219,220]. NFκB1 is activated by shear stress in endothelial cells and regulates various proinflammatory genes as a shear stress-sensitive transcriptional factor [221–224]. NFκB1 may thus link shear stress with chronic inflammation of aneurysm walls. Hypertension induced NFκB1 activation in the rat model of aortic aneurysm [225]. These observations suggest that NFκB1 is a key mediator of the formation of intracranial aneurysm [206].

Delayed cerebral ischemia is a common and serious complication following subarachnoid hemorrhage after ruptured intracranial aneurysm [226,227] (Fig. 12.10). Although this complication is at times reversible, it may develop into a cerebral infarction [229]. Delayed cerebral ischemia occurs in 20%–40% of patients and is associated with increased mortality and poor prognosis [230–232]. It is usually caused by a vasospasm [233], which remains a major cause of poor neurological outcome and increased mortality in the course of subarachnoid hemorrhage [230–232]. Vasospasm occurs between the 3rd and 15th day after the hemorrhage, with a peak at the 10th day. It is observed in 70% of patients and causes symptoms in 50% [233–235]. Vasospasm preferentially involves the vessels of the cranial base but also may affect small caliber vessels or diffusely the entire cerebral vascularization. The subsequent decrease in cerebral blood flow in the spastic arteries leads to delayed cerebral ischemia, which may develop into cerebral infarction [233,236].

The etiology of vasospasm is complex and still poorly understood. Several factors have been shown to be involved, such as endothelial dysfunction, loss of autoregulation, and a hypovolemic component leading to a decrease in cerebral blood flow [237,238]. At the acute phase the presence of oxyhemoglobin in the subarachnoid spaces causes a local and systemic inflammatory reaction with activation of platelets and coagulation factors [234,235,239]. The products derived from red blood cells, such as oxyhemoglobin, and from endothelial cells, such as endothelin 1 and reactive oxygen species, are considered to be mediators of the vasospasm [240,241]. The poor prognosis of patients with delayed cerebral ischemia following subarachnoid hemorrhage remains a major issue responsible for death and infirmity. Although our understanding of the physiopathology of delayed cerebral ischemia and vasospasm has improved, patient outcome has not been significantly modified [228].



**FIGURE 12.10** Proposed mechanisms of delayed cerebral ischemia after subarachnoid hemorrhage [228]. Delayed cerebral ischemia is usually caused by a vasospasm. Several factors, including endothelial dysfunction, loss of autoregulation, and a hypovolemic component leading to a decrease in cerebral blood flow, may be involved in the etiology of vasospasm. The products derived from red blood cells, such as oxyhemoglobin, and from endothelial cells, such as endothelin 1 and reactive oxygen species, are mediators of the vasospasm.

### 12.4.7 Molecular genetics of intracranial aneurysm and subarachnoid hemorrhage

Subarachnoid hemorrhage is 1.6 times more common in women than in men [191]. Hormonal factors probably explain this sex-specific risk, given that it is higher in postmenopausal women than in premenopausal women [242]. Smoking, alcohol consumption, and hypertension are also common risk factors for aneurysmal subarachnoid hemorrhage [243,244]. In addition to these environmental risk factors, genetic factors play an important role in the pathogenesis of subarachnoid hemorrhage associated with intracranial aneurysms. First-degree relatives of affected individuals are thus at up to seven times greater risk than the general population [245–249], and ~10% of patients with aneurysmal subarachnoid hemorrhage have first- or second-degree relatives with subarachnoid hemorrhage or unruptured intracranial aneurysms [246–248,250–252].

Several additional lines of evidence further support a role for genetic factors in the etiology of intracranial aneurysm. First, several genetic diseases, such as adult polycystic kidney disease [198], Marfan syndrome [253], glucocorticoid-remediable aldosteronism [254], and Ehlers–Danlos syndrome type IV [255], increase the risk of the formation of intracranial aneurysm. Second, familial recurrence of nonsyndromic intracranial aneurysm has been described [256–259]. Indeed, there is a three to fivefold increase in the risk for this condition in first-degree relatives of affected individuals compared with the general population [260,261]. Previous genetic linkage analyses have mapped loci for intracranial aneurysm to chromosomal regions 1p34.3–p36.13 [262], 2p13 [263], 5q22–q31 [264], 7q11 [264], 11q24–q25 [265], 14q22 [264], 14q23–q31 [265], 17cen [266], 19q13 [266], 19q13.3 [267], and Xp22 [266].

Magnetic resonance angiography (MRA) is not sufficiently effective for screening the first-degree relatives of individuals with sporadic subarachnoid hemorrhage for intracranial aneurysms [268], and repeated screening is necessary to identify newly developed aneurysms in familial subarachnoid hemorrhage [269]. Given that a familial predisposition is the strongest risk factor for the development of intracranial aneurysms [191,243], the identification of genetic risk factors might provide further diagnostic capability. In the future, genotype assessment might thus help to identify first-degree relatives of individuals with subarachnoid hemorrhage who are at high risk of developing one or more intracranial aneurysms [190]. Furthermore, identification of these genetic factors should provide insight into the pathophysiology of intracranial aneurysms [190]. The identification of disease susceptibility genes and increased understanding of the disease pathophysiology may lead to new therapeutic interventions to prevent the development, growth, or rupture of intracranial aneurysms [190].

Although various candidate gene association studies of unrelated individuals have identified genes that are related to subarachnoid hemorrhage or intracranial aneurysm (Table 12.4), the genes that confer susceptibility to these conditions have not been determined definitively. Recent GWASs identified several loci and genes that confer susceptibility for intracranial aneurysm (Table 12.5). Candidate genes for intracranial aneurysm or subarachnoid hemorrhage that are of particular interest [elastin gene (*ELN*), LIM domain kinase 1 gene (*LIMK1*), tumor necrosis factor receptor superfamily, member 13B gene (*TNFRSF13B*), and five loci identified by GWASs] are reviewed in the next section.

#### 12.4.7.1 Elastin gene and LIM domain kinase 1 gene

A functional haplotype spanning *ELN* and *LIMK1* was shown to confer susceptibility to intracranial aneurysm [273]. *ELN* is located within the chromosome 7q11 linkage region and was recognized earlier to be a positional and functional candidate gene for intracranial aneurysm [264]. However, allelic association studies yielded variable results [287,288]. A systematic analysis of 166 SNPs and haplotypes that reside within the chromosome 7q linkage peak identified a highly significant association between intracranial aneurysm and a distinct linkage disequilibrium block containing the 3′ untranslated region of *ELN* and the promoter region of *LIMK1* [273]. The strongest association was found with the *ELN* +695G→C tag SNP for a risk haplotype comprising the functional *ELN* +502A insertion and the *LIMK1* −187C→T SNP. Both the genotype and haplotype associations were replicated in an independent cohort. Functional studies revealed that the *ELN* +502A insertion reduces the rate of *ELN* transcription, whereas the *LIMK1* −187C→T SNP reduces promoter activity [189]. Synergism between genetic variants of *ELN* and *LIMK1* in their effects on vascular stability and distensibility seems plausible because (1) elastin is a major structural component of the internal elastic lamina in cerebral arteries; (2) *ELN* plays a key role in vascular development and remodeling; (3) secreted elastin activates a G protein–coupled signaling pathway that stimulates organization of actin stress fibers; and (4) *LIMK1* is a regulator of the actin cytoskeleton [189].

#### 12.4.7.2 Tumor necrosis factor receptor superfamily, member 13B gene

Sequence variation in *TNFRSF13B* was shown to contribute to risk for intracranial aneurysm [279]. Sequence analysis of genes in a linkage peak on chromosome 17p revealed several potentially deleterious changes in

**TABLE 12.4** Genes shown to be related to intracranial aneurysm or subarachnoid hemorrhage by linkage analyses or candidate gene association studies.

Chromosomal locus	Gene name	Gene symbol	References
1p36.1	Heparan sulfate proteoglycan 2	<i>HSPG2</i>	[270]
5q12-q14	Versican	<i>VCAN</i>	[271]
5q23-q31	Fibrillin 2	<i>FBN2</i>	[270]
6p21.3	Tumor necrosis factor	<i>TNF</i>	[80]
7p21	Interleukin 6	<i>IL6</i>	[272]
7q11.2	Elastin	<i>ELN</i>	[273]
7q11.23	Lim domain kinase 1	<i>LIMK1</i>	[273]
7q21.1	Cytochrome P450, family 3, subfamily A, polypeptide 4	<i>CYP3A4</i>	[76]
7q21.3-q22	Serpin peptidase inhibitor, clade E, member 1	<i>SERPINE1</i>	[270]
7q22.1	Collagen, type I, alpha 2	<i>COL1A2</i>	[274]
7q36	Nitric oxide synthase 3	<i>NOS3</i>	[275]
9q34.1	Endoglin	<i>ENG</i>	[276]
11q13	Uncoupling protein 3	<i>UCP3</i>	[80]
13q34	Collagen, type IV, alpha 1	<i>COL4A1</i>	[270]
14q32.1	Serpin peptidase inhibitor, clade A, member 3	<i>SERPINA3</i>	[277]
16p13.3-p13.12	Polycystic kidney disease 1	<i>PKD1</i>	[278]
17p11.2	Tumor necrosis factor receptor superfamily, member 13B	<i>TNFRSF13B</i>	[279]
17q21.32	Integrin, beta 3	<i>ITGB3</i>	[280]
17q23	Angiotensin I converting enzyme	<i>ACE</i>	[281]
20q11.2-q13.1	Matrix metalloproteinase 9	<i>MMP9</i>	[282]
22q12	Heme oxygenase 1	<i>HMOX1</i>	[283]

**TABLE 12.5** Chromosomal loci and genes shown to be related to intracranial aneurysm by genome-wide association studies.

Chromosomal locus	dbSNP	Nucleotide substitution	Gene (nearby gene)	References
2q33.1	rs700651	G → A	<i>BOLL</i>	[284]
4q31.22	rs6842241	A → C	<i>EDNRA</i>	[285]
4q31.22	rs6841581	G → A	<i>EDNRA</i>	[285]
8q11.12-12.1	rs10958409	G → A	<i>SOX17</i>	[284]
8q12.1	rs9298506	A → G	<i>SOX17</i>	[286]
9p21.3	rs1333040	C → T	<i>CDKN2B-AS1</i>	[284]
9p21.3	rs10757272	C → T	<i>CDKN2B-AS1</i>	[285]
10q24.32	rs12413409	G → A	<i>CNNM2</i>	[286]
13q13.1	rs9315204	C → T	<i>STARD13</i>	[286]
18q11.2	rs11661542	C → A	<i>RBBP8</i>	[286]



*TNFRSF13B* that segregated with intracranial aneurysm in pedigrees. Sequencing of a portion of *TNFRSF13B* in a large case-control sample showed that several potentially functional rare variants were more frequent in cases than in controls. Finally, association analyses suggested that one of the *TNFRSF13B* haplotypes was protective. Interactions of genetic factors such as *TNFRSF13B* with known risk factors for aneurysm formation, such as smoking and hypertension, remain an important area of research [189].

#### 12.4.7.3 Five loci for intracranial aneurysm identified by genome-wide association studies

A GWAS of intracranial aneurysm including Finnish, Dutch, and Japanese cohorts of 2196 cases and 8085 controls showed that common SNPs on 2q33.1, 8q11.12-12.1, and 9p21.3 were significantly associated with intracranial aneurysm [284]. In a follow-up GWAS, additional European cases and controls were included and the original Japanese replication cohort was increased, resulting in a cohort of 5891 cases and 14,181 controls [286]. This follow-up study identified three new loci strongly associated with intracranial aneurysms on 18q11.2, 13q13.1, and 10q24.32. The previously discovered associations of 8q11.23-q12.1 and 9p21.3 were replicated [286].

The 8q11.23-q12.1 locus contains the SRY-box 17 gene (*SOX17*), which encodes a transcription factor that has a pivotal role in endothelial cell function [289]. The strongly associated SNP was located in the 9p21.3 locus that contains *CDKN2B-AS1*. A recently described mutant mouse with a deletion corresponding to the human 9p21 locus showed a marked suppression of the expression of *CDKN2B* and *CDKN2A* [141]. Aortic smooth muscle cells in culture from these mice showed increased proliferative activity compared with aortic smooth muscle cells from wild-type mice [141]. The associated SNP on the 10q24.32 locus is located within the cyclin M2 gene (*CNNM2*). The 13q13.1 locus includes the StAR-related lipid transfer domain containing 13 gene (*STARD13*), of which overexpression leads to suppression in cell proliferation [290]. The 18q11.2 locus contains the retinoblastoma binding protein 8 gene (*RBBP8*). *RBBP8* is one of the proteins that bind directly to retinoblastoma protein, which regulates cell proliferation [291]. The gene products of the candidate genes in the five loci may be involved in cell proliferation [292].

## 12.5 Clinical implication

The increasing body of information garnered from studies on the genetics of stroke has resulted in the emergence of a greater understanding of the etiology of ischemic and hemorrhagic stroke. Such knowledge may have clinical implications for the prediction, diagnosis, prognosis, and treatment of stroke. The genes responsible for the pathogenesis of ischemic stroke, intracerebral hemorrhage, intracranial aneurysm, and subarachnoid hemorrhage as well as their encoded proteins are potentially important therapeutic targets in the design of new treatments for stroke. Genetic markers are potential diagnostic tools for the assessment of individuals at the risk of developing stroke. Genetic markers of stroke together with examination with MRI and MRA might also form the basis for promotion of preventive therapies in individuals at risk of stroke. It should be remembered, however, that gene–gene and gene–environment interactions make interpretation of information based on genetic markers of stroke more complex than is that of information based on markers for monogenic stroke. Another use of genetic markers might be to distinguish treatment responders from nonresponders (such as clopidogrel resistance) and to identify patients who are at risk of developing unfavorable side effects. It is likely that the use of gene polymorphisms to predict the response to adverse effects of therapies for stroke will increase in the future and will give rise to major advances in patient care. Genetic analysis of strokes is thus likely to have important direct clinical applications.

## 12.6 Conclusion

In this chapter, I have summarized single-gene disorders associated with stroke and the candidate loci and genes for common forms of stroke. There has been a growing effort to find genetic variants that confer risk for stroke as a means to understand the underlying biological events. Clarification of the functional relevance of SNPs at various loci to ischemic stroke, intracerebral hemorrhage, and subarachnoid hemorrhage may provide insight into the pathogenesis of these conditions as well as into the role of genetic factors in their development. Such studies may ultimately lead to the personalized prevention of stroke. It may thus become possible to predict the future risk for stroke in each individual on the basis of conventional laboratory analyses, examination with MRI and MRA, and genetic analyses. Furthermore, it may be possible to prevent an individual from having

stroke by medical intervention based on his or her genotypes for specific SNPs. In the future, we may have the ability to use specific agents particularized for certain genetic susceptibility factors, thereby increasing efficacy and limiting side effects of treatment. Identification of susceptibility genes for stroke and clarification of the functional relevance of genetic variants to this condition will thus contribute to the personalized prevention, early diagnosis, and treatment of stroke.

## References

- [1] Roger VL, Go AS, Lloyd-Jones DM, Benjamin EJ, Berry JD, Borden WB, et al. Heart disease and stroke statistics—2012 update: a report from the American Heart Association. *Circulation* 2012;125:e2–e220.
- [2] Warlow C, Sudlow C, Dennis M, Wardlaw J, Sandercock P. Stroke. *Lancet* 2003;362:1211–24.
- [3] Bak S, Gaist D, Sindrup SH, Skytthe A, Christensen K. Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke* 2002;33:769–74.
- [4] Meschia JF, Worrall BB. New advances in identifying genetic anomalies in stroke-prone probands. *Curr Atheroscler Rep* 2003;5:317–23.
- [5] Williams RR, Hunt SC, Heiss G, Province MA, Bensen JT, Higgins M, et al. Usefulness of cardiovascular family history data for population-based preventive medicine and medical research (the Health Family Tree Study and the NHLBI Family Heart Study). *Am J Cardiol* 2001;87:129–35.
- [6] Tournier-Lasserre E, Iba-Zizen MT, Romero N, Bousser MG. Autosomal dominant syndrome with strokelike episodes and leukoencephalopathy. *Stroke* 1991;22:1297–302.
- [7] Hassan A, Markus HS. Genetics and ischaemic stroke. *Brain* 2000;123:1784–812.
- [8] Hademenos GJ, Alberts MJ, Awad I, Mayberg M, Shepard T, Jagoda A, et al. Advances in the genetics of cerebrovascular disease and stroke. *Neurology* 2001;56:997–1008.
- [9] Natowicz M, Kelley RI. Mendelian etiologies of stroke. *Ann Neurol* 1987;22:175–92.
- [10] Tournier-Lasserre E. New players in the genetics of stroke. *N Engl J Med* 2002;347:1711–12.
- [11] Markus HS. Stroke genetics. *Hum Mol Genet* 2011;20(R2):R124–31.
- [12] Carr FJ, McBride MW, Carswell HV, Graham D, Strahorn P, Clark JS, et al. Genetic aspects of stroke: human and experimental studies. *J Cerebr Blood Flow Metab* 2002;22:767–73.
- [13] Tournier-Lasserre E, Joutel A, Melki J, Weissenbach J, Lathrop GM, Chabriat H, et al. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12. *Nat Genet* 1993;3:256–9.
- [14] Chabriat H, Vahedi K, Iba-Zizen MT, Joutel A, Nibbio A, Nagy TG, et al. Clinical spectrum of CADASIL: a study of 7 families. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. *Lancet* 1995;346:934–9.
- [15] Joutel A, Corpechot C, Ducros A, Vahedi K, Chabriat H, Mouton P, et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* 1996;383:707–10.
- [16] Joutel A, Corpechot C, Ducros A, Vahedi K, Chabriat H, Mouton P, et al. Notch3 mutations in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), a Mendelian condition causing stroke and vascular dementia. *Ann NY Acad Sci* 1997;826:213–17.
- [17] Opherk C, Peters N, Holtmannspötter M, Gschwendtner A, Müller-Myhsok B, Dichgans M. Heritability of MRI lesion volume in CADASIL: evidence for genetic modifiers. *Stroke* 2006;37:2684–9.
- [18] Hara K, Shiga A, Fukutake T, Nozaki H, Miyashita A, Yokoseki A, et al. Association of HTRA1 mutations and familial ischemic cerebral small-vessel disease. *N Engl J Med* 2009;360:1729–39.
- [19] Levy E, Carman MD, Fernandez-Madrid IJ, Power MD, Lieberburg I, van Duinen SG, et al. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science* 1990;248:1124–6.
- [20] Jansson O, Gudmundsson G, Arnason A, Blondal H, Petursdottir I, Thorsteinsson L, et al. Hereditary cystatin C (gamma-trace) amyloid angiopathy of the CNS causing cerebral hemorrhage. *Acta Neurol Scand* 1987;76:102–14.
- [21] Palsdottir A, Abrahamson M, Thorsteinsson L, Arnason A, Olafsson I, Grubb A, et al. Mutation in cystatin C gene causes hereditary brain haemorrhage. *Lancet* 1988;2:603–4.
- [22] Vidal R, Frangione B, Rostagno A, Mead S, Revesz T, Plant G, et al. A stop-codon mutation in the BRI gene associated with familial British dementia. *Nature* 1999;399:776–81.
- [23] Laberge-le Couteux S, Jung HH, Labauge P, Houtteville JP, Lescoat C, Cecillon M, et al. Truncating mutations in CCM1, encoding KRIT1, cause hereditary cavernous angiomas. *Nat Genet* 1999;23:189–93.
- [24] Richards A, van den Maagdenberg AM, Jen JC, Kavanagh D, Bertram P, Spitzer D, et al. C-terminal truncations in human 3'-5' DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy. *Nat Genet* 2007;39:1068–70.
- [25] Alamowitch S, Plaisier E, Favrole P, Prost C, Chen Z, Van Agtmael T, et al. Cerebrovascular disease related to COL4A1 mutations in HANAC syndrome. *Neurology* 2009;73:1873–82.
- [26] Meschia JF. Advances in genetics 2010. *Stroke* 2011;42:285–7.
- [27] Montagna P, Gallassi R, Medori R, Govoni E, Zeviani M, Di Mauro S, et al. MELAS syndrome: characteristic migrainous and epileptic features and maternal transmission. *Neurology* 1988;38:751–4.
- [28] Pavlakis SG, Phillips PC, DiMauro S, De Vivo DC, Rowland LP. Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes: a distinctive clinical syndrome. *Ann Neurol* 1984;16:481–8.
- [29] Enter C, Muller-Hocker J, Zierz S, Kurlmann G, Pongratz D, Forster C, et al. A specific point mutation in the mitochondrial genome of Caucasians with MELAS. *Hum Genet* 1991;88:233–6.
- [30] Sakuta R, Goto Y, Horai S, Nonaka I. Mitochondrial DNA mutations at nucleotide positions 3243 and 3271 in mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes: a comparative study. *J Neurol Sci* 1993;115:158–60.

- [31] Macmillan C, Lach B, Shoubridge EA. Variable distribution of mutant mitochondrial DNAs (tRNA(Leu[3243])) in tissues of symptomatic relatives with MELAS: the role of mitotic segregation. *Neurology* 1993;43:1586–90.
- [32] Ciafaloni E, Ricci E, Shanske S, Moraes CT, Silvestri G, Hirano M, et al. MELAS: clinical features, biochemistry, and molecular genetics. *Ann Neurol* 1992;31:391–8.
- [33] Pulkes T, Sweeney MG, Hanna MG. Increased risk of stroke in patients with the A12308G polymorphism in mitochondria. *Lancet* 2000;356:2068–9.
- [34] Spittell PC, Spittell Jr JA, Joyce JW, Tajik AJ, Edwards WD, Schaff HV, et al. Clinical features and differential diagnosis of aortic dissection: experience with 236 cases (1980 through 1990). *Mayo Clin Proc* 1993;68:642–51.
- [35] Schievink WI, Limburg M, Oorthuys JWE, Fleury P, Pope FM. Cerebrovascular disease in Ehlers–Danlos syndrome type IV. *Stroke* 1990;21:626–32.
- [36] Crutchfield KE, Patronas NJ, Dambrosia JM, Frei KP, Banerjee TK, Barton NW, et al. Quantitative analysis of cerebral vasculopathy in patients with Fabry disease. *Neurology* 1998;50:1746–9.
- [37] Floßmann E, Schulz UGR, Rothwell PM. Systematic review of methods and results of studies of the genetic epidemiology of ischemic stroke. *Stroke* 2004;35:212–27.
- [38] Gulcher JR, Gretarsdottir S, Helgadóttir A, Stefansson K. Genes contributing to risk for common forms of stroke. *Trends Mol Med* 2005;11:217–24.
- [39] Wilhelmsen L, Rosengren A, Eriksson H, Lappas G. Heart failure in the general population of men—morbidity, risk factors and prognosis. *J Intern Med* 2001;249:253–61.
- [40] Wilhelmsen L, Rosengren A, Lappas G. Hospitalizations for atrial fibrillation in the general male population: morbidity and risk factors. *J Intern Med* 2001;250:382–9.
- [41] Woodruff TM, Thundiyil J, Tang SC, Sobey CG, Taylor SM, Arumugam TV. Pathophysiology, treatment, and animal and cellular models of human ischemic stroke. *Mol Neurodegener* 2011;6:11.
- [42] Majno G, Joris I. Apoptosis, oncosis, and necrosis. An overview of cell death. *Am J Pathol* 1995;146:3–15.
- [43] Broughton BRS, Reutens DC, Sobey CG. Apoptotic mechanisms after cerebral Ischemia. *Stroke* 2009;40:e331–9.
- [44] Ginsberg MD. The new language of cerebral ischemia. *Am J Neuroradiol* 1997;18:1435–45.
- [45] Lipton P. Ischemic cell death in brain neurons. *Physiol Rev* 1999;79:1431–568.
- [46] Zheng Z, Yenari MA. Post-ischemic inflammation: molecular mechanisms and therapeutic implications. *Neurol Res* 2004;26:884–92.
- [47] Dirnagl U, Iadecola C, Moskowitz MA. Pathobiology of ischaemic stroke: an integrated view. *Trends Neurosci* 1999;22:391–7.
- [48] Mattson MP. Apoptosis in neurodegenerative disorders. *Nat Rev Mol Cell Biol* 2000;1:120–9.
- [49] Yu Z, Nikolova-Karakashian M, Zhou D, Cheng G, Schuchman EH, Mattson MP. Pivotal role for acidic sphingomyelinase in cerebral ischemia-induced ceramide and cytokine production, and neuronal death. *J Mol Neurosci* 2000;15:85–97.
- [50] Bonventre JV, Huang Z, Taheri MR, O’Leary E, Li E, Moskowitz MA, et al. Reduced fertility and postischemic brain injury in mice deficient in cytosolic phospholipase A2. *Nature* 1997;390:622–5.
- [51] Adams Jr HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* 1993;24:35–41.
- [52] Morgan L, Humphries SE. The genetics of stroke. *Curr Opin Lipidol* 2005;16:193–9.
- [53] Jerrard-Dunne P, Cloud G, Hassan A, Markus HS. Evaluating the genetic component of ischemic stroke subtypes: a family history study. *Stroke* 2003;34:1364–9.
- [54] Brass LM, Isaacssohn JL, Merikangas KR, Robinette CD. A study of twins and stroke. *Stroke* 1992;23:221–3.
- [55] Goldstein LB, Adams R, Becker K, Furberg CD, Gorelick PB, Hademenos G, et al. Primary prevention of ischemic stroke: a statement for healthcare professionals from the Stroke Council of the American Heart Association. *Stroke* 2001;32:280–99.
- [56] Humphries SE, Morgan L. Genetic risk factors for stroke and carotid atherosclerosis: insights into pathophysiology from candidate gene approaches. *Lancet Neurol* 2004;3:227–35.
- [57] Gretarsdottir S, Sveinbjornsdottir S, Jonsson HH, Jakobsson F, Einarsdottir E, Agnarsson U, et al. Localization of a susceptibility gene for common forms of stroke to 5q12. *Am J Hum Genet* 2002;70:593–603.
- [58] Morita H, Kurihara H, Tsubaki S, Sugiyama T, Hamada C, Kurihara Y, et al. Methylenetetrahydrofolate reductase gene polymorphism and ischemic stroke in Japanese. *Arterioscler Thromb Vasc Biol* 1998;18:1465–9.
- [59] Rubattu S, Stanzione R, Di Angelantonio E, Zanda B, Evangelista A, Tarasi D, et al. Atrial natriuretic peptide gene polymorphisms and risk of ischemic stroke in humans. *Stroke* 2004;35:814–18.
- [60] Morita A, Nakayama T, Soma M. Association study between C-reactive protein genes and ischemic stroke in Japanese subjects. *Am J Hypertens* 2006;19:593–600.
- [61] Zee RY, Cook NR, Cheng S, Reynolds R, Erlich HA, Lindpaintner K, et al. Polymorphism in the P-selectin and interleukin-4 genes as determinants of stroke: a population-based, prospective genetic analysis. *Hum Mol Genet* 2004;13:389–96.
- [62] Cipollone F, Toniato E, Martinotti S, Fazio M, Iezzi A, Cuccurullo C, et al. A polymorphism in the cyclooxygenase 2 gene as an inherited protective factor against myocardial infarction and stroke. *J Am Med Assoc* 2004;291:2221–8.
- [63] Iacoviello L, Di Castelnuovo A, Gattone M, Pezzini A, Assanelli D, Lorenzet R, et al. Polymorphisms of the interleukin-1 $\beta$  gene affect the risk of myocardial infarction and ischemic stroke at young age and the response of mononuclear cells to stimulation in vitro. *Arterioscler Thromb Vasc Biol* 2005;25:222–7.
- [64] Worrall BB, Brott TG, Brown Jr RD, Brown WM, Rich SS, Arepalli S, et al. IL1RN VNTR polymorphism in ischemic stroke. Analysis in 3 populations. *Stroke* 2007;38:1189–96.
- [65] Laverne E, Labreuche J, Daoudi M, Debre P, Cambien F, Deterre P, et al. Adverse associations between CX3CR1 polymorphisms and risk of cardiovascular or cerebrovascular disease. *Arterioscler Thromb Vasc Biol* 2005;25:847–53.
- [66] Lee BC, Lee HJ, Chung JH. Peroxisome proliferator-activated receptor- $\gamma$ 2 Pro12Ala polymorphism is associated with reduced risk for ischemic stroke with type 2 diabetes. *Neurosci Lett* 2006;410:141–5.

- [67] Morrison AC, Doris PA, Folsom AR, Nieto FJ, Boerwinkle E, Atherosclerosis Risk in Communities Study. G-protein  $\beta 3$  subunit and  $\alpha$ -adducin polymorphisms and risk of subclinical and clinical stroke. *Stroke* 2001;32:822–9.
- [68] Kessler C, Spitzer C, Stauske D, Mende S, Stadlmuller J, Walther R, et al. The apolipoprotein E and  $\beta$ -fibrinogen G/A–455 gene polymorphisms are associated with ischemic stroke involving large-vessel disease. *Arterioscler Thromb Vasc Biol* 1997;17:2880–4.
- [69] Carlsson M, Orho-Melander M, Hedenbro J, Almgren P, Groop LC. The T 54 allele of the intestinal fatty acid-binding protein 2 is associated with a parental history of stroke. *J Clin Endocrinol Metab* 2000;85:2801–4.
- [70] Gretarsdottir S, Thorleifsson G, Reynisdottir ST, Manolescu A, Jonsdottir S, Jonsdottir T, et al. The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. *Nat Genet* 2003;35:131–8.
- [71] Carlsson LE, Santoso S, Spitzer C, Kessler C, Greinacher A. The  $\alpha_2\beta_1$  gene coding sequence T807/A873 of the platelet collagen receptor integrin  $\alpha_2\beta_1$  might be a genetic risk factor for the development of stroke in younger patients. *Blood* 1999;93:3583–6.
- [72] Voetsch B, Jin RC, Bierl C, Benke KS, Kenet G, Simioni P, et al. Promoter polymorphisms in the plasma glutathione peroxidase (GPx-3) gene: a novel risk factor for arterial ischemic stroke among young adults and children. *Stroke* 2007;38:41–9.
- [73] Santamaria A, Mateo J, Tirado I, Oliver A, Belvis R, Marti-Fabregas J, et al. Homozygosity of the T allele of the 46 C→T polymorphism in the F12 gene is a risk factor for ischemic stroke in the Spanish population. *Stroke* 2004;35:1795–9.
- [74] Elbaz A, Poirier O, Canaple S, Chedru F, Cambien F, Amarenco P. The association between the Val34Leu polymorphism in the factor XIII gene and brain infarction. *Blood* 2000;95:586–91.
- [75] Szolnoki Z, Havasi V, Talian G, Bene J, Komlosi K, Somogyvari F, et al. Lymphotoxin- $\alpha$  gene 252G allelic variant is a risk factor for large-vessel-associated ischemic stroke. *J Mol Neurosci* 2005;27:205–11.
- [76] Yamada Y, Metoki N, Yoshida H, Satoh K, Kato K, Hibino T, et al. Genetic factors for ischemic and hemorrhagic stroke in Japanese individuals. *Stroke* 2008;39:2211–18.
- [77] Shearman AM, Cooper JA, Kotwinski PJ, Humphries SE, Mendelsohn ME, Housman DE, et al. Estrogen receptor  $\alpha$  gene variation and the risk of stroke. *Stroke* 2005;36:2281–2.
- [78] Sun L, Li Z, Zhang H, Ma A, Liao Y, Wang D, et al. Pentanucleotide TTTTA repeat polymorphism of apolipoprotein(a) gene and plasma lipoprotein(a) are associated with ischemic and hemorrhagic stroke in Chinese: a multicenter case-control study in China. *Stroke* 2003;34:1617–22.
- [79] Pola R, Flex A, Gaetani E, Flore R, Serricchio M, Pola P. Synergistic effect of –174 G/C polymorphism of the interleukin-6 gene promoter and 469 E/K polymorphism of the intercellular adhesion molecule-1 gene in Italian patients with history of ischemic stroke. *Stroke* 2003;34:881–5.
- [80] Yamada Y, Metoki N, Yoshida H, Satoh K, Ichihara S, Kato K, et al. Genetic risk for ischemic and hemorrhagic stroke. *Arterioscler Thromb Vasc Biol* 2006;26:1920–5.
- [81] Voetsch B, Benke KS, Damasceno BP, Siqueira LH, Loscalzo J. Paraoxonase 192 Gln→Arg polymorphism: an independent risk factor for nonfatal arterial ischemic stroke among young adults. *Stroke* 2002;33:1459–64.
- [82] Wiklund PG, Nilsson L, Ardnor SN, Eriksson P, Johansson L, Stegmayr B, et al. Plasminogen activator inhibitor-1 4G/5G polymorphism and risk of stroke: replicated findings in two nested case-control studies based on independent cohorts. *Stroke* 2005;36:1661–5.
- [83] Elbaz A, Poirier O, Moulin T, Chedru F, Cambien F, Amarenco P. Association between the Glu298Asp polymorphism in the endothelial constitutive nitric oxide synthase gene and brain infarction. The GENIC Investigators. *Stroke* 2000;31:1634–9.
- [84] Shimo-Nakanishi Y, Urabe T, Hattori N, Watanabe Y, Nagao T, Yokochi M, et al. Polymorphism of the lipoprotein lipase gene and risk of atherothrombotic cerebral infarction in the Japanese. *Stroke* 2001;32:1481–6.
- [85] Fornage M, Lee CR, Doris PA, Bray MS, Heiss G, Zeldin DC, et al. The soluble epoxide hydrolase gene harbors sequence variation associated with susceptibility to and protection from incident ischemic stroke. *Hum Mol Genet* 2005;14:2829–37.
- [86] Saito K, Nakayama T, Sato N, Morita A, Takahashi T, Soma M, et al. Haplotypes of the plasminogen activator gene associated with ischemic stroke. *Thromb Haemost* 2006;96:331–6.
- [87] Anderson CD, Biffi A, Rost NS, Cortellini L, Furie KL, Rosand J. Chromosome 9p21 in ischemic stroke: population structure and meta-analysis. *Stroke* 2010;41:1123–31.
- [88] De Stefano V, Chiusolo P, Paciaroni K, Casorelli I, Rossi E, Molinari M, et al. Prothrombin G20210A mutant genotype is a risk factor for cerebrovascular ischemic disease in young patients. *Blood* 1998;91:3562–5.
- [89] Havasi V, Szolnoki Z, Talian G, Bene J, Komlosi K, Maasz A, et al. Apolipoprotein A5 gene promoter region T–1131C polymorphism associates with elevated circulating triglyceride levels and confers susceptibility for development of ischemic stroke. *J Mol Neurosci* 2006;29:177–83.
- [90] Hsieh K, Lalouschek W, Schillinger M, Endler G, Reisinger M, Janisw M, et al. Impact of  $\alpha$ ENaC polymorphisms on the risk of ischemic cerebrovascular events: a multicenter case-control study. *Clin Chem* 2005;51:952–6.
- [91] Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, Thorsteinsdottir U, et al. The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. *Nat Genet* 2004;36:233–9.
- [92] Herrmann SM, Funke-Kaiser H, Schmidt-Petersen K, Nicaud V, Gautier-Bertrand M, Evans A, et al. Characterization of polymorphic structure of cathepsin G gene: role in cardiovascular and cerebrovascular diseases. *Arterioscler Thromb Vasc Biol* 2001;21:1538–43.
- [93] Hegener HH, Diehl KA, Kurth T, Gaziano JM, Ridker PM, Zee RY. Polymorphisms of prostaglandin-endoperoxide synthase 2 gene, and prostaglandin-E receptor 2 gene, C-reactive protein concentrations and risk of atherothrombosis: a nested case-control approach. *J Thromb Haemost* 2006;4:1718–22.
- [94] Wang Y, Zhang W, Zhang Y, Yang Y, Sun L, Hu S, et al. VKORC1 haplotypes are associated with arterial vascular diseases (stroke, coronary heart disease, and aortic dissection). *Circulation* 2006;113:1615–21.
- [95] Ito D, Murata M, Watanabe K, Yoshida T, Saito I, Tanahashi N, et al. C242T polymorphism of NADPH oxidase p22 PHOX gene and ischemic cerebrovascular disease in the Japanese population. *Stroke* 2000;31:936–9.
- [96] Baker RI, Eikelboom J, Lofthouse E, Staples N, Afshar-Kharghan V, Lopez JA, et al. Platelet glycoprotein Iba Kozak polymorphism is associated with an increased risk of ischemic stroke. *Blood* 2001;98:36–40.



- [97] Ridker PM, Hennekens CH, Schmitz C, Stampfer MJ, Lindpaintner K. PIA1/A2 polymorphism of platelet glycoprotein IIIa and risks of myocardial infarction, stroke, and venous thrombosis. *Lancet* 1997;349:385–8.
- [98] Margaglione M, Celentano E, Grandone E, Vecchione G, Cappucci G, Giuliani N, et al. Deletion polymorphism in the angiotensin-converting enzyme gene in patients with a history of ischemic stroke. *Arterioscler Thromb Vasc Biol* 1996;16:304–9.
- [99] Kaneko Y, Nakayama T, Saito K, Morita A, Sato I, Maruyama A, et al. Relationship between the thromboxane A2 receptor gene and susceptibility to cerebral infarction. *Hypertens Res* 2006;29:665–71.
- [100] Frikke-Schmidt R, Nordestgaard BG, Schnohr P, Tybjaerg-Hansen A. Single nucleotide polymorphism in the low-density lipoprotein receptor is associated with a threefold risk of stroke. A case-control and prospective study. *Eur Heart J* 2004;25:943–51.
- [101] Kim Y, Lee C. The gene encoding transforming growth factor  $\beta 1$  confers risk of ischemic stroke and vascular dementia. *Stroke* 2006;37:2843–5.
- [102] Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol* 2008;64:402–9.
- [103] The International Stroke Genetics Consortium (ISGC), the Wellcome Trust Case Control Consortium (WTCCC), Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet* 2012;44:328–33.
- [104] Hata J, Matsuda K, Ninomiya T, Yonemoto K, Matsushita T, Ohnishi Y, et al. Functional SNP in an Sp1-binding site of AGTRL1 gene is associated with susceptibility to brain infarction. *Hum Mol Genet* 2007;16:630–9.
- [105] Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med* 2009;360:1718–28.
- [106] Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, et al. A nonsynonymous SNP in PRKCH (protein kinase C-eta) increases the risk of cerebral infarction. *Nat Genet* 2007;39:212–17.
- [107] Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, et al. A sequence variant in ZFXH3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet* 2009;41:876–8.
- [108] DeBette S, Bis JC, Fornage M, Schmidt H, Ikram MA, Sigurdsson S, et al. Genome-wide association studies of MRI-defined brain infarcts: meta-analysis from the CHARGE consortium. *Stroke* 2010;41:210–17.
- [109] Yamada Y, Fuku N, Tanaka M, Aoyagi Y, Sawabe M, Metoki N, et al. Identification of CELSR1 as a susceptibility gene for ischemic stroke in Japanese individuals by a genome-wide association study. *Atherosclerosis* 2009;207:144–9.
- [110] Fukumoto S, Koyama H, Hosoi M, Yamakawa K, Tanaka S, Morii H, et al. Distinct role of cAMP and cGMP in the cell cycle control of vascular smooth muscle cells: cGMP delays cell cycle transition through suppression of cyclin D1 and cyclin-dependent kinase 4 activation. *Circ Res* 1999;85:985–91.
- [111] Palmer D, Tsoi K, Maurice DH. Synergistic inhibition of vascular smooth muscle cell migration by phosphodiesterase 3 and phosphodiesterase 4 inhibitors. *Circ Res* 1998;82:852–61.
- [112] Pan X, Arauz E, Krzanowski JJ, Fitzpatrick DF, Polson JB. Synergistic interactions between selective pharmacological inhibitors of phosphodiesterase isozyme families PDE III and PDE IV to attenuate proliferation of rat vascular smooth muscle cells. *Biochem Pharmacol* 1994;48:827–35.
- [113] Houslay MD, Adams DR. PDE4 cAMP phosphodiesterases: modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization. *Biochem J* 2003;370:1–18.
- [114] Indolfi C, Avvedimento EV, Di Lorenzo E, Esposito G, Rapacciuolo A, Giuliano P, et al. Activation of cAMP–PKA signaling *in vivo* inhibits smooth muscle cell proliferation induced by vascular injury. *Nat Med* 1997;3:775–9.
- [115] Indolfi C, Di Lorenzo E, Rapacciuolo A, Stingone AM, Stabile E, Leccia A, et al. 8-Chloro-cAMP inhibits smooth muscle cell proliferation *in vitro* and neointima formation induced by balloon injury *in vivo*. *J Am Coll Cardiol* 2000;36:288–93.
- [116] Libby P. Inflammation in atherosclerosis. *Nature* 2002;420:868–74.
- [117] Lusis AJ. Atherosclerosis. *Nature* 2000;407:233–41.
- [118] Naghavi M, Libby P, Falk E, Casscells SW, Litovsky S, Rumberger J, et al. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: part I. *Circulation* 2003;108:1664–72.
- [119] Markus HS, Alberts MJ. Update on genetics of stroke and cerebrovascular disease 2005. *Stroke* 2006;37:288–90.
- [120] Bevan S, Porteous L, Sitzer M, Markus HS. Phosphodiesterase 4D gene, ischemic stroke, and asymptomatic carotid atherosclerosis. *Stroke* 2005;36:949–53.
- [121] Meschia JF, Brott TG, Brown Jr RD, Crook R, Worrall BB, Kissela B, et al. Phosphodiesterase 4D and 5-lipoxygenase activating protein in ischemic stroke. *Ann Neurol* 2005;58:351–61.
- [122] Lohmussaar E, Gschwendtner A, Mueller JC, Org T, Wichmann E, Hamann G, et al. ALOX5AP gene and the PDE4D gene in a central European population of stroke patients. *Stroke* 2005;36:731–6.
- [123] Nilsson-Ardnor S, Wiklund PG, Lindgren P, Nilsson AK, Janunger T, Escher SA, et al. Linkage of ischemic stroke to the PDE4D region on 5q in a Swedish population. *Stroke* 2005;36:1666–71.
- [124] Bevan S, Dichgans M, Gschwendtner A, Kühlenbäumer G, Ringelstein EB, Markus HS. Variation in the PDE4D gene and ischemic stroke risk. A systematic review and meta-analysis on 5200 cases and 6600 controls. *Stroke* 2008;39:1966–71.
- [125] Newcombe PJ, Verzilli C, Casas JP, Hingorani AD, Smeeth L, Whittaker JC. Multilocus Bayesian meta-analysis of gene-disease associations. *Am J Hum Genet* 2009;84:567–80.
- [126] Helgadottir A, Gretarsdottir S, Clair S, D, Manolescu A, Cheung J, Thorleifsson G, et al. Association between the gene encoding 5-lipoxygenase-activating protein and stroke replicated in a Scottish population. *Am J Hum Genet* 2005;76:505–9.
- [127] Dixon RA, Diehl RE, Opas E, Rands E, Vickers PJ, Evans JF, et al. Requirement of a 5-lipoxygenase-activating protein for leukotriene synthesis. *Nature* 1990;343:282–4.
- [128] Spanbroek R, Grabner R, Lotzer K, Hildner M, Urbach A, Ruhling K, et al. Expanding expression of the 5-lipoxygenase pathway within the arterial wall during human atherogenesis. *Proc Natl Acad Sci USA* 2003;100:1238–43.

- [129] Mehrabian M, Allayee H, Wong J, Shi W, Wang XP, Shaposhnik Z, et al. Identification of 5-lipoxygenase as a major gene contributing to atherosclerosis susceptibility in mice. *Circ Res* 2002;91:120–6.
- [130] Zintzaras E, Rodopoulou P, Sakellaridis N. Variants of the arachidonate 5-lipoxygenase-activating protein (ALOX5AP) gene and risk of stroke: a HuGE gene-disease association review and meta-analysis. *Am J Epidemiol* 2009;169:523–32.
- [131] McPherson R, Pertsemliadis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science* 2007;316:1488–91.
- [132] Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* 2007;357:443–53.
- [133] Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–78.
- [134] Helgadóttir A, Thorleifsson G, Manolescu A, Gretarsdóttir S, Blondal T, Jonasdóttir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 2007;316:1491–3.
- [135] Matarin M, Brown WM, Singleton A, Hardy JA, Meschia JF, ISGS investigators. Whole genome analyses suggest ischemic stroke and heart disease share an association with polymorphisms on chromosome 9p21. *Stroke* 2008;39:1586–9.
- [136] Gschwendtner A, Bevan S, Cole JW, Plourde A, Matarin M, Ross-Adams H, et al. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. *Ann Neurol* 2009;65:531–9.
- [137] Helgadóttir A, Thorleifsson G, Magnusson KP, Grétarsdóttir S, Steinthorsdóttir V, Manolescu A, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* 2008;40:217–24.
- [138] Musunuru K, Post WS, Herzog W, Shen H, O'Connell JR, McArdle PF, et al. Association of single nucleotide polymorphisms on chromosome 9p21.3 with platelet reactivity: a potential mechanism for increased vascular disease. *Circ Cardiovasc Genet* 2010;3:445–53.
- [139] Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, Naing T, et al. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler Thromb Vasc Biol* 2009;29:1671–7.
- [140] Harismendy O, Notani D, Song X, Rahim NG, Tanasa B, Heintzman N, et al. 9p21 DNA variants associated with coronary artery disease impair interferon- $\gamma$  signalling response. *Nature* 2011;470:264–8.
- [141] Visel A, Zhu Y, May D, Afzal V, Gong E, Attanasio C, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. *Nature* 2010;464:409–12.
- [142] Cole JW, Meschia JF. Stroke genetics update: 2011. *Curr Cardiovasc Risk Rep* 2011;5:533–41.
- [143] Rosand J, Meschia JF, Singleton AB. Failure to validate association between 12p13 variants and ischemic stroke. *N Engl J Med* 2010;362:1547–50.
- [144] Lotta LA, Giusti B, Saracini C, Vestri A, Volpe M, Rubattu S, et al. No association between chromosome 12p13 single nucleotide polymorphisms and early-onset ischemic stroke. *J Thromb Haemost* 2010;8:1858–60.
- [145] Olsson S, Melander O, Jood K, Smith JG, Lövkvist H, Sjögren M, et al. Genetic variant on chromosome 12p13 does not show association to ischemic stroke in 3 Swedish case-control studies. *Stroke* 2011;42:214–16.
- [146] Lemmens R, Buyschaert I, Geelen V, Fernandez-Cadenas I, Montaner J, Schmidt H, et al. The association of the 4q25 susceptibility variant for atrial fibrillation with stroke is limited to stroke of cardioembolic etiology. *Stroke* 2010;41:1850–7.
- [147] Qureshi AI, Tuhim S, Broderick JP, Batjer HH, Hondo H, Hanley DF. Spontaneous intracerebral hemorrhage. *N Engl J Med* 2001;344:1450–60.
- [148] Cole FM, Yates PO. Pseudo-aneurysms in relationship to massive cerebral hemorrhage. *J Neurol Neurosurg Psychiatry* 1967;30:61–6.
- [149] Qureshi AI, Mendelow AD, Hanley DF. Intracerebral haemorrhage. *Lancet* 2009;373:1632–44.
- [150] Wagner KR, Sharp FR, Ardizzone TD, Lu A, Clark JF. Heme and iron metabolism: role in cerebral hemorrhage. *J Cereb Blood Flow Metab* 2003;23:629–52.
- [151] Xi G, Keep RF, Hoff JT. Mechanisms of brain injury after intracerebral haemorrhage. *Lancet Neurol* 2006;5:53–63.
- [152] Ardizzone TD, Lu A, Wagner KR, Tang Y, Ran R, Sharp FR. Glutamate receptor blockade attenuates glucose hypermetabolism in perihematoma brain after experimental intracerebral hemorrhage in rat. *Stroke* 2004;35:2587–91.
- [153] Qureshi AI, Ali Z, Suri MF, Shuaib A, Baker G, Todd K, et al. Extracellular glutamate and other amino acids in experimental intracerebral hemorrhage: an in vivo microdialysis study. *Crit Care Med* 2003;31:1482–9.
- [154] Mun-Bryce S, Wilkerson AC, Papuashvili N, Okada YC. Recurring episodes of spreading depression are spontaneously elicited by an intracerebral hemorrhage in the swine. *Brain Res* 2001;888:248–55.
- [155] Aronowski J, Hall CE. New horizons for primary intracerebral hemorrhage treatment: experience from preclinical studies. *Neurol Res* 2005;27:268–79.
- [156] Tang J, Liu J, Zhou C, Ostanin D, Grisham MB, Neil Granger D, et al. Role of NADPH oxidase in the brain injury of intracerebral hemorrhage. *J Neurochem* 2005;94:1342–50.
- [157] Wang J, Doré S. Inflammation after intracerebral hemorrhage. *J Cereb Blood Flow Metab* 2007;27:894–908.
- [158] Wang J, Fields J, Zhao C, Langer J, Thimmulappa RK, Kensler TW, et al. Role of Nrf2 in protection against intracerebral hemorrhage injury in mice. *Free Radic Biol Med* 2007;43:408–14.
- [159] Xue M, Del Bigio MR. Intracerebral injection of autologous whole blood in rats: time course of inflammation and cell death. *Neurosci Lett* 2000;283:230–2.
- [160] Gong C, Hoff JT, Keep RF. Acute inflammatory reaction following experimental intracerebral hemorrhage in rat. *Brain Res* 2000;871:57–65.
- [161] Hickenbottom SL, Grotta JC, Strong R, Denner LA, Aronowski J. Nuclear factor-kappa B and cell death after experimental intracerebral hemorrhage in rats. *Stroke* 1999;30:2472–7.
- [162] Aronowski J, Zhao X. Molecular pathophysiology of cerebral hemorrhage: secondary brain injury. *Stroke* 2011;42:1781–6.
- [163] Huang FP, Xi G, Keep RF, Hua Y, Nemoianu A, Hoff JT. Brain edema after experimental intracerebral hemorrhage: role of hemoglobin degradation products. *J Neurosurg* 2002;96:287–93.



- [164] Wang X, Mori T, Sumii T, Lo EH. Hemoglobin-induced cytotoxicity in rat cerebral cortical neurons: caspase activation and oxidative stress. *Stroke* 2002;33:1882–8.
- [165] Felberg RA, Grotta JC, Shirzadi AL, Strong R, Narayana P, Hill-Felberg SJ, et al. Cell death in experimental intracerebral hemorrhage: the “black hole” model of hemorrhagic damage. *Ann Neurol* 2002;51:517–24.
- [166] Zhao X, Sun G, Zhang J, Strong R, Song W, Gonzales N, et al. Hematoma resolution as a target for intracerebral hemorrhage treatment: role for peroxisome proliferator-activated receptor gamma in microglia/macrophages. *Ann Neurol* 2007;61:352–62.
- [167] Zhao X, Zhang Y, Strong R, Zhang J, Grotta JC, Aronowski J. Distinct patterns of intracerebral hemorrhage-induced alterations in NF-kappaB subunit, iNOS, and COX-2 expression. *J Neurochem* 2007;101:652–63.
- [168] Masada T, Hua Y, Xi G, Yang GY, Hoff JT, Keep RF. Attenuation of intracerebral hemorrhage and thrombin-induced brain edema by overexpression of interleukin-1 receptor antagonist. *J Neurosurg* 2001;95:680–6.
- [169] Mayne M, Ni W, Yan HJ, Xue M, Johnston JB, Del Bigio MR, et al. Antisense oligodeoxynucleotide inhibition of tumor necrosis factor-alpha expression is neuroprotective after intracerebral hemorrhage. *Stroke* 2001;32:240–8.
- [170] Rosenberg GA. Matrix metalloproteinases in neuroinflammation. *Glia* 2002;39:279–91.
- [171] Tang J, Liu J, Zhou C, Alexander JS, Nanda A, Granger DN, et al. MMP-9 deficiency enhances collagenase-induced intracerebral hemorrhage and brain injury in mutant mice. *J Cereb Blood Flow Metab* 2004;24:1133–45.
- [172] Alberts MJ, McCarron MO, Hoffmann KL, Graffagnino C. Familial clustering of intracerebral hemorrhage: a prospective study in North Carolina. *Neuroepidemiology* 2002;21:18–21.
- [173] Massaro AR, Sacco RL, Mohr JP, Foulkes MA, Tatemichi TK, Price TR, et al. Clinical discriminators of lobar and deep hemorrhages: the Stroke Data Bank. *Neurology* 1991;41:1881–5.
- [174] Sacco RL. Lobar intracerebral hemorrhage. *N Engl J Med* 2000;342:276–9.
- [175] O'Donnell HC, Rosand J, Knudsen KA, Furie KL, Segal AZ, Chiu RI, et al. Apolipoprotein E genotype and the risk of recurrent lobar intracerebral hemorrhage. *N Engl J Med* 2000;342:240–5.
- [176] Woo D, Kaushal R, Chakraborty R, Woo J, Haverbusch M, Sekar P, et al. Association of apolipoprotein E4 and haplotypes of the apolipoprotein E gene with lobar intracerebral hemorrhage. *Stroke* 2005;36:1874–9.
- [177] Greenberg SM, Briggs ME, Hyman BT, Kokoris GJ, Takis C, Kanter DS, et al. Apolipoprotein E epsilon 4 is associated with the presence and earlier onset of hemorrhage in cerebral amyloid angiopathy. *Stroke* 1996;27:1333–7.
- [178] Reiner AP, Schwartz SM, Frank MB, Longstreth Jr WT, Hindorff LA, Teramura G, et al. Polymorphisms of coagulation factor XIII subunit A and risk of nonfatal hemorrhagic stroke in young white women. *Stroke* 2001;32:2580–6.
- [179] Alberts MJ, Davis JP, Graffagnino C, McClenny C, Delong D, Granger C, et al. Endoglin gene polymorphism as a risk factor for sporadic intracerebral hemorrhage. *Ann Neurol* 1997;41:683–6.
- [180] Gould DB, Phalan FC, van Mil SE, Sundberg JP, Vahedi K, Massin P, et al. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *N Engl J Med* 2006;354:1489–96.
- [181] Vila N, Obach V, Revilla M, Oliva R, Chamorro A.  $\alpha_1$ -Antichymotrypsin gene polymorphism in patients with stroke. *Stroke* 2000;31:2103–5.
- [182] Slowik A, Turaj W, Dziedzic T, Haefele A, Pera J, Malecki MT, et al. DD genotype of ACE gene is a risk factor for intracerebral hemorrhage. *Neurology* 2004;63:359–61.
- [183] Xia J, Yang QD, Yang QM, Xu HW, Liu YH, Zhang L, et al. Apolipoprotein H gene polymorphisms and risk of primary cerebral hemorrhage in a Chinese population. *Cerebrovasc Dis* 2004;17:197–203.
- [184] Nicoll JA, Burnett C, Love S, Graham DI, Dewar D, Ironside JW, et al. High frequency of apolipoprotein E epsilon 2 allele in hemorrhage due to cerebral amyloid angiopathy. *Ann Neurol* 1997;41:716–21.
- [185] Greenberg SM, Rebeck GW, Vonsattel JP, Gomez-Isla T, Hyman BT. Apolipoprotein E epsilon 4 and cerebral hemorrhage associated with amyloid angiopathy. *Ann Neurol* 1995;38:254–9.
- [186] Greenberg SM, Vonsattel JP, Segal AZ, Chiu RI, Clatworthy AE, Liao A, et al. Association of apolipoprotein E epsilon 2 and vasculopathy in cerebral amyloid angiopathy. *Neurology* 1998;50:961–5.
- [187] Gould DB, Phalan FC, Breedveld GJ, van Mil SE, Smith RS, Schimenti JC, et al. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. *Science* 2005;308:1167–71.
- [188] van der Knaap MS, Smit LM, Barkhof F, Pijnenburg YA, Zweegman S, Niessen HW, et al. Neonatal porencephaly and adult stroke related to mutations in collagen IV A1. *Ann Neurol* 2006;59:504–11.
- [189] Dichgans M, Hegele RA. Update on the genetics of stroke and cerebrovascular disease 2006. *Stroke* 2007;38:216–18.
- [190] Ruigrok YM, Rinkel GJ, Wijmenga C. Genetics of intracranial aneurysms. *Lancet Neurol* 2005;4:179–89.
- [191] Rinkel GJE, Djibuti M, Algra A, van Gijn J. Prevalence and risk of rupture of intracranial aneurysms. *Stroke* 1998;29:251–6.
- [192] Inagawa T, Tokuda Y, Ohbayashi N, Takaya M, Moritake K. Study of aneurysmal subarachnoid hemorrhage in Izumo City, Japan. *Stroke* 1995;26:761–6.
- [193] Longstreth WT, Nelson LM, Koepsell TD, van Belle G. Clinical course of spontaneous subarachnoid hemorrhage: a population-based study in King County, Washington. *Neurology* 1993;43:712–18.
- [194] Hop JW, Rinkel GJE, Algra A, van Gijn J. Case fatality rates and functional outcome after subarachnoid haemorrhage: a systematic review. *Stroke* 1997;28:660–4.
- [195] Linn FHH, Rinkel GJE, Algra A, van Gijn J. Incidence of subarachnoid hemorrhage: role of region, year and rate of computed tomography: a meta-analysis. *Stroke* 1996;27:625–9.
- [196] Johnston SC, Selvin S, Gress DR. The burden, trends, and demographics of mortality from subarachnoid hemorrhage. *Neurology* 1998;50:1413–18.
- [197] Schievink WI, Michels VV, Piepgras DG. Neurovascular manifestations of heritable connective tissue disorders: a review. *Stroke* 1994;25:889–903.
- [198] Chapman AB, Rubinstein D, Hughes R, Stears JC, Earnest MP, Johnson AM, et al. Intracranial aneurysms in autosomal dominant polycystic kidney disease. *N Engl J Med* 1992;327:916–20.

- [199] Neil-Dwyer G, Bartlett JR, Nicholls AC, Narcisi P, Pope FM. Collagen deficiency and ruptured cerebral aneurysms: a clinical and biochemical study. *J Neurosurg* 1983;59:16–20.
- [200] Ostergaard JR, Oxlund H. Collagen type III deficiency in patients with rupture of intracranial saccular aneurysms. *J Neurosurg* 1987;67:690–6.
- [201] Ostergaard JR, Reske-Nielsen E, Oxlund H. Histological and morphometric observations on the reticular fibers in the arterial beds of patients with ruptured intracranial saccular aneurysms. *Neurosurgery* 1987;20:554–8.
- [202] Skirgaudas M, Awad IA, Kim J, Rothbart D, Criscuolo G. Expression of angiogenesis factors and selected vascular wall matrix proteins in intracranial saccular aneurysms. *Neurosurgery* 1996;39:537–45.
- [203] van den Berg JS, Limburg M, Pals G, Arwert F, Westerveld A, Hennekam RC, et al. Some patients with intracranial aneurysms have a reduced type III/type I collagen ratio. A case-control study. *Neurology* 1997;49:1546–51.
- [204] Chyatte D, Reilly J, Tilson MD. Morphometric analysis of reticular and elastin fibers in the cerebral arteries of patients with intracranial aneurysms. *Neurosurgery* 1990;26:939–43.
- [205] Hegedus K. Some observations on reticular fibers in the media of the major cerebral arteries: a comparative study of patients without vascular diseases and those with ruptured berry aneurysms. *Surg Neurol* 1984;22:301–7.
- [206] Aoki T, Nishimura M. Molecular mechanism of cerebral aneurysm formation focusing on NF- $\kappa$ B as a key mediator of inflammation. *J Biochem* 2010;24:16–21.
- [207] Chyatte D, Bruno G, Desai S, Todor DR. Inflammation and intracranial aneurysms. *Neurosurgery* 1999;45:1137–46.
- [208] Aoki T, Kataoka H, Morimoto M, Nozaki K, Hashimoto N. Macrophage-derived matrix metalloproteinase-2 and -9 promote the progression of cerebral aneurysms in rats. *Stroke* 2007;38:162–9.
- [209] Aoki T, Kataoka H, Ishibashi R, Nozaki K, Egashira K, Hashimoto N. Impact of monocyte chemoattractant protein-1 deficiency on cerebral aneurysm formation. *Stroke* 2009;40:942–51.
- [210] Gaetani P, Tartara F, Grazioli V, Tancioni F, Infuso L, Baena R. Collagen cross-linkage, elastolytic and collagenolytic activities in cerebral aneurysms: a preliminary investigation. *Life Sci* 1998;63:285–92.
- [211] Gaetani P, Tartara F, Tancioni F, Baena R, Casari E, Alfano M, et al. Deficiency of total collagen content and of deoxypyridinoline in intracranial aneurysm walls. *FEBS Lett* 1997;404:303–6.
- [212] Bruno G, Todor R, Lewis I, Chyatte D. Vascular extracellular matrix remodeling in cerebral aneurysms. *J Neurosurg* 1998;89:431–40.
- [213] Chyatte D, Lewis I. Gelatinase activity and the occurrence of cerebral aneurysms. *Stroke* 1997;28:799–804.
- [214] Kondo S, Hashimoto N, Kikuchi H, Hazama F, Nagata I, Kataoka H. Apoptosis of medial smooth muscle cells in the development of saccular cerebral aneurysms in rats. *Stroke* 1998;29:181–8.
- [215] Moriawaki T, Takagi Y, Sadamasa N, Aoki T, Nozaki K, Hashimoto N. Impaired progression of cerebral aneurysms in interleukin-1 $\beta$ -deficient mice. *Stroke* 2006;37:900–5.
- [216] Sadamasa N, Nozaki K, Hashimoto N. Disruption of gene for inducible nitric oxide synthase reduces progression of cerebral aneurysms. *Stroke* 2003;34:2980–4.
- [217] Fukuda S, Hashimoto N, Naritomi H, Nagata I, Nozaki K, Kondo S, et al. Prevention of rat cerebral aneurysm formation by inhibition of nitric oxide synthase. *Circulation* 2000;101:2532–8.
- [218] Guo F, Li Z, Song L, Han T, Feng Q, Guo Y, et al. Increased apoptosis and cysteinyl aspartate specific protease-3 gene expression in human intracranial aneurysm. *J Clin Neurosci* 2007;14:550–5.
- [219] Lianxu C, Hongti J, Changlong Y. NF- $\kappa$ Bp65-specific siRNA inhibits expression of genes of COX-2, NOS-2 and MMP-9 in rat IL-1 $\beta$ -induced and TNF- $\alpha$ -induced chondrocytes. *Osteoarthritis Cartil* 2006;14:367–76.
- [220] Ping D, Boekhoudt GH, Rogers EM, Boss JM. Nuclear factor- $\kappa$ B p65 mediates the assembly and activation of the TNF-responsive element of the murine monocyte chemoattractant-1 gene. *J Immunol* 1999;162:727–34.
- [221] Khachigian LM, Resnick N, Gimbrone Jr MA, Collins T. Nuclear factor- $\kappa$ B interacts functionally with the platelet-derived growth factor B-chain shear-stress response element in vascular endothelial cells exposed to fluid shear stress. *J Clin Invest* 1995;96:1169–75.
- [222] Ballermann BJ, Dardik A, Eng E, Liu A. Shear stress and the endothelium. *Kidney Int* 1998;67(Suppl):S100–8.
- [223] Orr AW, Sanders JM, Bevard M, Coleman E, Sarembock IJ, Schwartz MA. The subendothelial extracellular matrix modulates NF- $\kappa$ B activation by flow: a potential role in atherosclerosis. *J Cell Biol* 2005;169:191–202.
- [224] Davis ME, Grumbach IM, Fukai T, Cutchins A, Harrison DG. Shear stress regulates endothelial nitric-oxide synthase promoter activity through nuclear factor  $\kappa$ B binding. *J Biol Chem* 2004;279:163–8.
- [225] Shiraya S, Miwa K, Aoki M, Miyake T, Oishi M, Kataoka K, et al. Hypertension accelerated experimental abdominal aortic aneurysm through upregulation of nuclear factor  $\kappa$ B and Ets. *Hypertension* 2006;48:628–36.
- [226] Pluta RM, Hansen-Schwartz J, Dreier J, Vajkoczy P, Macdonald RL, Nishizawa S, et al. Cerebral vasospasm following subarachnoid hemorrhage: time for a new world of thought. *Neurol Res* 2009;31:151–8.
- [227] de Rooij NK, Linn FH, van der Plas JA, Algra A, Rinkel GJ. Incidence of subarachnoid haemorrhage: a systematic review with emphasis on region, age, gender and time trends. *J Neurol Neurosurg Psychiatry* 2007;78:1365–72.
- [228] Castanares-Zapatero D, Hantson P. Pharmacological treatment of delayed cerebral ischemia and vasospasm in subarachnoid hemorrhage. *Ann Intensive Care* 2011;1:12.
- [229] Fergusen S, Macdonald RL. Predictors of cerebral infarction in patients with aneurysmal subarachnoid hemorrhage. *Neurosurgery* 2007;60:658–67.
- [230] Macdonald RL, Pluta RM, Zhang JH. Cerebral vasospasm after subarachnoid hemorrhage: the emerging revolution. *Nat Clin Pract Neurol* 2007;3:256–63.
- [231] Rabinstein AA, Friedman JA, Weigand SD, McClelland RL, Fulgham JR, Manno EM, et al. Predictors of cerebral infarction in aneurysmal subarachnoid hemorrhage. *Stroke* 2004;35:1862–6.
- [232] Dorsch NW, King MT. A review of cerebral vasospasm in aneurysmal subarachnoid haemorrhage Part I: Incidence and effects. *J Clin Neurosci* 1994;1:19–26.

- [233] Vajkoczy P, Horn P, Thome C, Munch E, Schmiedek P. Regional cerebral blood flow monitoring in the diagnosis of delayed ischemia following aneurysmal subarachnoid hemorrhage. *J Neurosurg* 2003;98:1227–34.
- [234] Kassell NF, Sasaki T, Colohan AR, Nazar G. Cerebral vasospasm following aneurysmal subarachnoid hemorrhage. *Stroke* 1985;16:562–72.
- [235] Heros RC, Zervas NT, Varsos V. Cerebral vasospasm after subarachnoid hemorrhage: an update. *Ann Neurol* 1983;14:599–608.
- [236] Suarez JJ, Tarr RW, Selman WR. Aneurysmal subarachnoid hemorrhage. *N Engl J Med* 2006;354:387–96.
- [237] Diringier MN. Subarachnoid hemorrhage: a multiple-organ system disease. *Crit Care Med* 2003;31:1884–5.
- [238] Dumont AS, Dumont RJ, Chow MM, Lin CL, Calisaneller T, Ley KF, et al. Cerebral vasospasm after subarachnoid hemorrhage: putative role of inflammation. *Neurosurgery* 2003;53:123–33.
- [239] Yoshimoto Y, Tanaka Y, Hoya K. Acute systemic inflammatory response syndrome in subarachnoid hemorrhage. *Stroke* 2001;32:1989–93.
- [240] Macdonald RL, Weir BK. A review of hemoglobin and the pathogenesis of cerebral vasospasm. *Stroke* 1991;22:971–82.
- [241] Rubanyi GM. Endothelium-derived relaxing and contracting factors. *J Cell Biochem* 1991;46:27–36.
- [242] Longstreth WT, Nelson LM, Koepsell TD, van Belle G. Subarachnoid hemorrhage and hormonal factors in women. A population-based case-control study. *Ann Intern Med* 1994;121:168–73.
- [243] Ruigrok YM, Buskens E, Rinkel GJE. Attributable risk of common and rare determinants of subarachnoid hemorrhage. *Stroke* 2001;32:1173–5.
- [244] Teunissen LL, Rinkel GJE, Algra A, van Gijn J. Risk factors for subarachnoid hemorrhage: a systematic review. *Stroke* 1996;27:544–9.
- [245] Ronkainen A, Hernesniemi J, Puranen M, Niemitukia L, Vanninen R, Ryyanen M, et al. Familial intracranial aneurysms. *Lancet* 1997;349:380–4.
- [246] Braekeleer DM, Pêrussee L, Cantin L, Bouchard JM, Mathieu J. A study of inbreeding and kinship in intracranial aneurysms in the Saguenay Lac-Saint-Jean region (Quebec, Canada). *Ann Hum Genet* 1996;60:99–104.
- [247] Schievink WI, Schaid DJ, Michels VV, Piepgras DG. Familial aneurysmal subarachnoid hemorrhage: a community based study. *J Neurosurg* 1995;83:426–9.
- [248] Bromberg JE, Rinkel GJ, Algra A, Greebe P, van Duyn CM, Hasan D, et al. Subarachnoid haemorrhage in first and second degree relatives of patients with subarachnoid haemorrhage. *Br Med J* 1995;311:288–9.
- [249] Gaist D, Vaeth M, Tsiropoulos I, Christensen K, Corder E, Olsen J, et al. Risk of subarachnoid haemorrhage in first degree relatives of patients with subarachnoid haemorrhage: follow up study based on national registries in Denmark. *Brit Med J* 2000;320:141–5.
- [250] Norrgard O, Angquist KA, Fodstad H, Forsell A, Lindberg M. Intracranial aneurysms and heredity. *Neurosurgery* 1987;20:236–9.
- [251] Ronkainen A, Hernesniemi J, Ryyanen M. Familial subarachnoid hemorrhage in east Finland, 1977–1990. *Neurosurgery* 1993;33:787–96.
- [252] Wang PS, Longstreth Jr WT, Koepsell TD. Subarachnoid hemorrhage and family history: a population-based case-control study. *Arch Neurol* 1995;52:202–4.
- [253] ter Berg HW, Bijlsma JB, Veiga Pires JA, Ludwig JW, van der Heiden C, Tulleken CA, et al. Familial association of intracranial aneurysms and multiple congenital anomalies. *Arch Neurol* 1986;43:30–3.
- [254] Litchfield WR, Anderson BF, Weiss RJ, Lifton RP, Dluhy RG. Intracranial aneurysm and hemorrhagic stroke in glucocorticoid-remediable aldosteronism. *Hypertension* 1998;31:445–50.
- [255] de Paepe A, van Landegem W, de Keyser F, de Reuck J. Association of multiple intracranial aneurysms and collagen type III deficiency. *Clin Neurol Neurosurg* 1988;90:53–6.
- [256] Maroun FB, Murray GP, Jacob JC, Mangan MA, Faridi M. Familial intracranial aneurysms: report of three families. *Surg Neurol* 1986;25:85–8.
- [257] Morooka Y, Waga S. Familial intracranial aneurysms: report of four families. *Surg Neurol* 1983;19:260–2.
- [258] Elshunnar KS, Whittle IR. Familial intracranial aneurysms: report of five families. *Br J Neurosurg* 1990;4:181–6.
- [259] Fox JL, Ko JP. Familial intracranial aneurysms. Six cases among 13 siblings. *J Neurosurg* 1980;52:501–3.
- [260] Ronkainen A, Niskanen M, Piironen R, Hernesniemi J. Familial subarachnoid hemorrhage. Outcome study. *Stroke* 1999;30:1099–102.
- [261] Stehbens WE. Familial intracranial aneurysms: an autopsy study. *Neurosurgery* 1998;43:1258–9.
- [262] Nahed BV, Seker A, Guclu B, Ozturk AK, Finberg K, Hawkins AA, et al. Mapping a Mendelian form of intracranial aneurysm to 1p34.3–p36.13. *Am J Hum Genet* 2005;76:172–9.
- [263] Roos YB, Pals G, Struycken PM, Rinkel GJ, Limburg M, Pronk JC, et al. Genome-wide linkage in a large Dutch consanguineous family maps a locus for intracranial aneurysms to chromosome 2p13. *Stroke* 2004;35:2276–81.
- [264] Onda H, Kasuya H, Yoneyama T, Takakura K, Hori T, Takeda J, et al. Genomewide-linkage and haplotype-association studies map intracranial aneurysm to chromosome 7q11. *Am J Hum Genet* 2001;69:804–19.
- [265] Ozturk AK, Nahed BV, Bydon M, Bilguvar K, Goksu E, Bademci G, et al. Molecular genetic analysis of two large kindreds with intracranial aneurysms demonstrates linkage to 11q24-25 and 14q23-31. *Stroke* 2006;37:1021–7.
- [266] Yamada S, Utsunomiya M, Inoue K, Nozaki K, Inoue S, Takenaka K, et al. Genome-wide scan for Japanese familial intracranial aneurysms: linkage to several chromosomal regions. *Circulation* 2004;14:3727–33.
- [267] Van der Voet M, Olson JM, Kuivaniemi H, Dudek DM, Skunca M, Ronkainen A, et al. Intracranial aneurysms in Finnish families: confirmation of linkage and refinement of the interval to chromosome 19q13.3. *Am J Hum Genet* 2004;74:564–71.
- [268] The Magnetic Resonance Angiography in Relatives of Patients with Subarachnoid Hemorrhage Study Group. Risks and benefits of screening of intracranial aneurysms in first-degree relatives of patients with sporadic subarachnoid hemorrhage. *N Engl J Med* 1999;341:1344–50.
- [269] Raaymakers TWM, Rinkel GJE, Ramos LMP. Initial and follow up screening for aneurysms in familial subarachnoid hemorrhage. *Neurology* 1998;51:1125–30.
- [270] Ruigrok YM, Rinkel GJ, van't Slot R, Wolfs M, Tang S, Wijmenga C. Evidence in favor of the contribution of genes involved in the maintenance of the extracellular matrix of the arterial wall to the development of intracranial aneurysms. *Hum Mol Genet* 2006;15:3361–8.

- [271] Ruigrok YM, Rinkel GJ, Wijmenga C. The versican gene and the risk of intracranial aneurysms. *Stroke* 2006;37:2372–4.
- [272] Morgan L, Cooper J, Montgomery H, Kitchen N, Humphries SE. The interleukin-6 gene –174G→C and –572G→C promoter polymorphisms are related to cerebral aneurysms. *J Neurol Neurosurg Psychiatry* 2006;77:915–17.
- [273] Akagawa H, Tajima A, Sakamoto Y, Krischek B, Yoneyama T, Kasuya H, et al. A haplotype spanning two genes, ELN and LIMK1, decreases their transcripts and confers susceptibility to intracranial aneurysms. *Hum Mol Genet* 2006;15:1722–34.
- [274] Yoneyama T, Kasuya H, Onda H, Akagawa H, Hashiguchi K, Nakajima T, et al. Collagen type I  $\alpha 2$  (COL1A2) is the susceptible gene for intracranial aneurysms. *Stroke* 2004;35:443–8.
- [275] Khurana VG, Sohni YR, Mangrum WI, McClelland RL, O’Kane DJ, Meyer FB, et al. Endothelial nitric oxide synthase gene polymorphisms predict susceptibility to aneurysmal subarachnoid hemorrhage and cerebral vasospasm. *J Cerebr Blood Flow Metab* 2004;24:291–7.
- [276] Takenaka K, Sakai H, Yamakawa H, Yoshimura S, Kumagai M, Yamakawa H, et al. Polymorphism of the endoglin gene in patients with intracranial saccular aneurysms. *J Neurosurg* 1999;90:935–8.
- [277] Slowik A, Borratynska A, Turaj W, Pera J, Dziedzic T, Figlewicz DA, et al.  $\alpha_1$ -Antichymotrypsin gene (SERPINA3) A/T polymorphism as a risk factor for aneurysmal subarachnoid hemorrhage. *Stroke* 2005;36:737–40.
- [278] Watnick T, Phakdeekitcharoen B, Johnson A, Gandolph M, Wang M, Briefel G, et al. Mutation detection of PKD1 identifies a novel mutation common to three families with aneurysms and/or very-early-onset disease. *Am J Hum Genet* 1999;65:1561–71.
- [279] Inoue K, Mineharu Y, Inoue S, Yamada S, Matsuda F, Nozaki K, et al. Search on chromosome 17 centromere reveals TNFRSF13B as a susceptibility gene for intracranial aneurysm: a preliminary study. *Circulation* 2006;113:2002–10.
- [280] Iniesta JA, Gonzalez-Conejero R, Piqueras C, Vicente V, Corral J. Platelet GP IIIa polymorphism HPA-1 (PIA) protects against subarachnoid hemorrhage. *Stroke* 2004;35:2282–6.
- [281] Slowik A, Borratynska A, Pera J, Betlej M, Dziedzic T, Krzyszkowski T, et al. II Genotype of the angiotensin-converting enzyme gene increases the risk for subarachnoid hemorrhage from ruptured aneurysm. *Stroke* 2004;35:1594–7.
- [282] Peters DG, Kassam A, Jean St PL, Yonas H, Ferrell RE. Functional polymorphism in the matrix metalloproteinase-9 promoter as a potential risk factor for intracranial aneurysm. *Stroke* 1999;30:2612–16.
- [283] Morgan L, Hawe E, Palmen J, Montgomery H, Humphries SE, Kitchen N. Polymorphism of the heme oxygenase-1 gene and cerebral aneurysms. *Br J Neurosurg* 2005;19:317–21.
- [284] Bilguvar K, Yasuno K, Niemelä M, Ruigrok YM, von Und Zu Fraunberg M, van Duijn CM, et al. Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat Genet* 2008;40:1472–7.
- [285] Low SK, Takahashi A, Cha PC, Zembutsu H, Kamatani N, Kubo M, et al. Genome-wide association study for intracranial aneurysm in Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. *Hum Mol Genet* 2012;21(9):2102–10.
- [286] Yasuno K, Bilguvar K, Bijlenga P, Low SK, Krischek B, Auburger G, et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat Genet* 2010;42:420–5.
- [287] Ruigrok YM, Seitz U, Wolterink S, Rinkel GJ, Wijmenga C, Urban Z. Association of polymorphisms and haplotypes in the elastin gene in Dutch patients with sporadic aneurysmal subarachnoid hemorrhage. *Stroke* 2004;35:2064–8.
- [288] Hofer A, Hermans M, Kubassek N, Sitzler M, Funke H, Stogbauer F, et al. Elastin polymorphism haplotype and intracranial aneurysms are not associated in Central Europe. *Stroke* 2003;34:1207–11.
- [289] Francois M, Koopman P, Beltrame M. SoxF genes: key players in the development of the cardio-vascular system. *Int J Biochem Cell Biol* 2010;42:445–8.
- [290] Leung TH, Ching YP, Yam JW, Wong CM, Yau TO, Jin DY, et al. Deleted in liver cancer 2 (DLC2) suppresses cell transformation by means of inhibition of RhoA activity. *Proc Natl Acad Sci USA* 2005;102:15207–12.
- [291] Nevins JR. Toward an understanding of the functional complexity of the E2F and retinoblastoma families. *Cell Growth Differ* 1998;9:585–93.
- [292] Ruigrok YM, Rinkel GJ. From GWAS to the clinic: risk factors for intracranial aneurysms. *Genome Med* 2010;2:61.

## Further reading

Alberts, 2003    Alberts MJ. Stroke genetics update. *Stroke* 2003;34:342–4.

# Clinical molecular endocrinology

*Atul Kalhan*

Diabetes & Endocrinology, Royal Glamorgan Hospital, Llantrisant, United Kingdom

## 13.1 Introduction

Clinical endocrinology incorporates the diagnosis and management of disorders associated with hormone synthesis and/or function. Recent advances in the ever so expanding field of molecular biology have provided clinicians with a novel insight into molecular mechanism leading to endocrine dysfunction. This chapter on clinical molecular endocrinology brings together commonly encountered scenarios from clinical practice in endocrinology with emphasis on molecular mechanisms that underpin the clinical presentation. The clinical cases that have been discussed in detail cover relevant molecular and genetic pathways leading to pituitary, parathyroid, adrenal, and gonadal dysfunction.

## 13.2 Congenital hypopituitarism, congenital hypogonadotropic hypogonadism, and pituitary adenoma

### 13.2.1 Congenital hypopituitarism

#### 13.2.1.1 Introduction

Congenital hypopituitarism (CH) is a relatively rare condition characterized by deficiency of one or multiple anterior pituitary hormones. It is mostly sporadic although there are some rare familial forms. The etiology remains unknown in most cases although over 30 genes have been implicated in pathogenesis of CH [1]. Neonates and children with CH may have isolated or combined deficiency of growth hormone (GH), adrenocorticotrophic hormone (ACTH), thyrotropin-stimulating hormone (TSH), luteinizing hormone (LH), follicle-stimulating hormone (FSH), and prolactin. GH deficiency remains the commonest hormone abnormality seen in majority of the patients with CH [2]. In view of separate molecular pathways involved in pathogenesis of congenital hypogonadotropic hypogonadism (CHH), it has been described separately in Section 13.2.2.

##### 13.2.1.1.1 Clinical case

A male neonate was noticed to be lethargic and in respiratory distress within 24 hours of a full-term normal vaginal delivery from a 31-year-old gravida 2, para 1 woman. The pregnancy as well as intrapartum period had been uncomplicated. The mother had no past medical history of any systemic illness and apart from folic acid tablets, which she was prescribed during pregnancy, she was not on any regular medications. There was no family history of pituitary or endocrine disorders.

On examination, the neonate had a respiratory rate of 60/minute, although pulse rate, blood pressure (BP), and body temperature were within normal range. Oxygen saturation was 94% in room air. There were no obvious morphological skeletal or facial abnormalities. The systemic examination was unremarkable.



## Investigations:

White cell count	$9 \times 10^3/\mu\text{L}$	
C-reactive protein	2.0 mg/dL	(<5)
Blood glucose	1.1 mmol/L	
FT4	0.5 pmol/L	(10–36)
TSH	0.8 mU/L	(1.1–6.3)
IGF1	Not detectable	
Short Synacthen test:		
0-hour cortisol	<25 nmol/L	
30-minute cortisol	84 nmol/L	
Chest X-ray	Clear lung fields; normal-sized heart	
Magnetic resonance imaging	Normal pituitary gland; no sellar/suprasellar masses	

This neonate had presented with features of respiratory distress within hours of birth although there were no clinical or biochemical features suggestive of any infective etiology. He had severe hypoglycemia on initial biochemical investigations, which prompted further anterior hypopituitary hormone screening. He had multiple anterior pituitary hormone deficiencies (GH, ACTH, and TSH) consistent with a diagnosis of CH.

The signs and symptoms associated with CH are nonspecific and depend upon underlying hormone defect and age of onset. In neonatal period, presenting features of CH include lethargy, failure to thrive, prolonged cholestatic jaundice, hypoglycemia, seizures, and respiratory distress. Boys with CH may have micropenis or cryptorchidism. Hypoglycemia in such patients is indicative of ACTH and/or GH deficiency.

Children with CH may present with clinical features of an isolated hormone deficiency (GH deficiency being the commonest) although results from a long-term observational study suggest that such patients may go on to develop multiple hormonal deficiencies at a later stage [3].

CH may manifest as an isolated pituitary defect or be part of a characteristic syndrome with midline craniofacial defects such as

- septo-optic dysplasia characterized by optic nerve hypoplasia, agenesis of corpus callosum, absence of septum pellucidum (in 60%) [4],
- holoprosencephaly, and
- short cervical spine.

**Overview of the relevant molecular systems underpinning the clinical scenario** Mutations in over 30 genes involved in pituitary ontogenesis and function have been implicated in pathogenesis of CH. The fully penetrant genes are detected only in 20% of the patients with familial CH while the frequency of known gene mutations in sporadic cases of CH remains even lower [5].

In the given clinical scenario, there is clinical and biochemical evidence of multiple pituitary hormone deficiencies (ACTH, TSH, and GH) although there are no midline craniofacial defects or family history of hypopituitarism. The most common genes that are involved in isolated CH include *Prop1* and *POU1F1*.

The genetic and clinical characteristics of CH subtypes is shown in Table 13.1 [6–8].

**Management of congenital hypopituitarism** Management of suspected patients with CH includes a careful evaluation of anterior pituitary hormone function and radiological imaging (magnetic resonance imaging) of the pituitary gland to assess for pituitary hypoplasia and to exclude structural sellar/extrasellar lesions. Patients may require dynamic pituitary tests, such as Growth Hormone Releasing Hormone (GHRH)–arginine stimulation test, clonidine stimulation test, or insulin tolerance test, to confirm the initial anterior-pituitary hormone-screening results. Patients with CH require lifelong hormone replacement therapy and surveillance.



**TABLE 13.1** The genetic and clinical characteristics of congenital hypopituitarism (CH) subtypes.

Condition	Gene	Hormone deficiency	Radiological features
CH	<i>Prop1</i>	GH, TSH, gonadotropins, prolactin	Intracranial pseudotumor (in 40%) Normal stalk and posterior pituitary
Combined pituitary hormone deficiency	<i>POU1F1</i>	GH, prolactin, TSH	—
Kallmann syndrome	<i>KAL1, FGFR1, PROKR2, PROK2</i>	GnRH	Absent olfactory bulbs
Isolated ACTH deficiency	<i>Tpit</i>	ACTH	—
Septo-optic dysplasia	<i>HESX4</i>	Multiple pituitary hormones	Absent septum pellucidum and corpus callosum Optic nerve hypoplasia
Holoprosencephaly	<i>GLI2</i>	Multiple pituitary hormones	Severe cranial–facial defects
Miscellaneous	<i>LHX3, LHX4, OTX2</i>	Multiple pituitary hormones	<i>LHX3</i> : limited neck rotation <i>LHX4</i> : poorly developed sella turcica <i>OTX2</i> : microphthalmia (Stieg M)

ACTH, Adrenocorticotropic hormone; CH, congenital hypopituitarism; GH, growth hormone; GnRH, gonadotropin-releasing hormone; TSH, thyrotropin-stimulating hormone.

### 13.2.2 Congenital hypogonadotropic hypogonadism

#### 13.2.2.1 Introduction

CHH is a heterogeneous disorder resulting from deficient (or absent) secretion and/or function of gonadotropin-releasing hormone (GnRH) [9]. Although it was historically considered as a monogenic disorder, advances in translational research over the last decade have identified over 25 genetic mutations leading to CHH [10]. It is characterized clinically, be a delay or absence of puberty, along with infertility in affected individuals. A constitutional delay in growth and puberty (CDGP) remains the main differential diagnosis. There is a 3–4:1 male to female preponderance with heterogeneity in the mode of inheritance (X-linked, autosomal-recessive, or autosomal-dominant) [11].

#### 13.2.2.2 Clinical case

An 18-year-old boy was referred to an endocrinology clinic with features of delayed puberty. He had attained developmental milestones appropriately after birth and had no history of any systemic illness, weight loss, malabsorption, or eating disorder. He had noticed a lack of facial hair growth and his body hair remained sparse. He was otherwise in good health and was not on any medications. He had a 13-year-old brother who was going through normal pubertal changes. His sense of smell was intact.

On examination, his height was 172 cm and he had a body weight of 65 kg (BMI 21.9 kg/m<sup>2</sup>). He had bilateral gynecomastia. There were sparse facial and body hair growth. His testes were 2 mL in volume.

Investigations:

FSH	0.5 U/L	(1.4–18.1 U/L)
LH	0.8 U/L	(3.0–8.0 U/L)
Prolactin	250 mU/L	(45–375 mU/L)
Testosterone	1.3 nmol/L	(8.4–28.7 nmol/L)
Free T <sub>4</sub>	14.5 pmol/L	(11.5–22.7 pmol/L)
TSH	2.1 mU/L	(0.35–5.5 mU/L)

**TABLE 13.2** Correlation of underlying gene mutation with phenotype in patients with congenital hypogonadotropic hypogonadism.

Gene mutation	Phenotype
<i>KAL1</i>	Anosmia/hyposmia and synkinesis
<i>SOX10</i> or <i>IL17RD</i>	Anosmia/hyposmia and hearing loss
<i>CHD7</i>	Hearing loss, genital/urinary defects, coloboma, heart defects
<i>FGFR1</i>	Cleft lip/palate, skeletal abnormalities
<i>LEP</i> , <i>LEPR</i> , <i>PCSK1</i>	Morbid obesity

### 13.2.2.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

This 18-year-old boy has clinical features of delayed puberty and absence of secondary sexual characteristics. Typically, onset of puberty in boys is seen at the age of 9–14 years (median age 11.5 years) and characterized by growth of testes size, development of axillary and pubic hair, increase in growth velocity, deepening of voice, and increased penile length. An absence of testicular enlargement (volume <4 mL) in boys aged >14 years is suggestive of delayed puberty. In girls, absence of thelarche at an age >13 years or primary amenorrhea at an age >15 years indicates delayed puberty [12,13].

The hypothalamic–pituitary–gonadal (HPG) axis plays a central role in initiation and development of puberty. The GnRH comprises decapeptides secreted by hypothalamic nuclei in response to peripheral metabolic cues such as steroid hormones and leptin. The pulsatile GnRH secretion heralds the onset of puberty. GnRH, in turn, acts on GnRH receptors located in anterior pituitary gland to stimulate the secretion of luteinizing hormone (LH) and follicular stimulatory hormone (FSH). LH acts on Leydig cells (or theca cells in ovaries) in testes to stimulate testosterone secretion [14], while FSH acts on Sertoli cells (or granulosa cells in ovaries) to induce spermatogenesis (ovarian follicular growth).

CHH is characterized by delayed or absent puberty and infertility due to deficiency in secretion or action of GnRH as is the case of the 18-year-old boy in the current clinical scenario. He has clinical (delayed puberty and lack of secondary sexual–characteristics development) as well as biochemical features (low testosterone levels along with inappropriately low LH and FSH levels) consistent with a diagnosis of CHH. CHH may be associated with other congenital anomalies such as cleft palate, dental agenesis, renal agenesis, and bimanual synkinesis. Kallmann syndrome is a variant of CHH resulting from a failure of embryonic migration of GnRH neurons from the olfactory placode to the forebrain. Early clues toward diagnosis of CHH include cryptorchidism (mal descended testes due to a lack of neonatal activation of HPG axis) in male infants and micropenis (a sign of GnRH deficiency). In contrast, genital abnormalities, such as hypospadias, are suggestive of human chorionic gonadotropin (hCG)–driven androgen secretion/action during embryonic development. Differential diagnosis of CHH includes CDGP, functional causes, such as anorexia/systemic illness, celiac disease, excessive exercise, systemic conditions, such as sarcoidosis, histiocytosis, thalassemia, hemochromatosis, and acquired etiologies such as pituitary tumor, Rathke cleft cyst, and medications such as opiates/steroids [15]. A lack of epiphyseal fusion results in an increased length of axial bones leading to eunuchoid habitus. In contrast, young adults with CDGP have short stature and delay in skeletal maturation. As patients with CHH may have associated phenotypic characteristics (Table 13.2), clinical examination should specifically seek presence or absence of abnormalities with sense of smell, skeletal abnormalities, such as cleft lip or palate, hearing loss, and synkinesis [15].

The baseline investigations for a patient with suspected CHH include serum testosterone (or estradiol in females), gonadotropins (FSH/LH), prolactin, thyroid function test, 9 a.m. cortisol, and IGF1 levels (to assess anterior pituitary function). Inhibin B, Anti Mullerian Hormone (AMH), and kisspeptin levels may be obtained, if available, as these serve as important biomarkers regarding puberty and have prognostic value. Radiological investigations, such as magnetic resonance imaging (MRI) of pituitary gland, can exclude structural lesions in addition to providing supportive evidence for lack of development of olfactory tract in patients with Kallmann syndrome. An ultrasound of pelvis may be helpful in boys with cryptorchidism to identify an inguinal or abdominal testis and in girls to evaluate size and texture of uterus and fallopian tubes. Bone-density assessment (DXA scan) and wrist radiography (bone age) may be useful in distinguishing CHH from CDGP.

**Overview of the relevant molecular systems underpinning the clinical scenario** GnRH is a decapeptide released into hypophyseal portal circulation, which acts via GnRH receptors on gonadotrophs in the anterior pituitary gland. GnRH-secreting neurons are specialized neuroendocrine cells with ectodermal and neural crest origin.

These migrate during embryonic life along with axons of olfactory nerve to the hypothalamus from their origin in the olfactory placode. Proteins, such as anosmin, fibroblast growth factor (FGF) 8, prokineticin 2, play a key role in migration of GnRH-secreting neurons. Mutations in the aforementioned genes can lead to the impaired migration of GnRH to its final abode near hypothalamus and result in CHH [10].

The following genes have been implicated in CHH:

FGF8

(Sex determining region Y)-box 10 (SOX10)

Anosmin 1 (formerly known as KAL1: Kallmann syndrome protein)

Prokineticin 2 (PROK2)

Kisspeptin 1 (KISS1) and kisspeptin receptor (KISS1R)

Tachykinin family (TAC3 and TACR3)

Kisspeptin and tachykinin family (TAC3 and TACR3)—producing neurons are afferents to GnRH neurons. Inactivating mutations in the kisspeptin gene and its receptor has been associated with an absence of puberty. Kisspeptin is a potent stimulator of GnRH-induced LH secretion.

Interestingly, the GnRH secretion may indirectly be modulated by metabolic signals from periphery. Inactivation of leptin or leptin receptor gene (LEP and LEPR) is associated with delayed or absent puberty. Kisspeptin neurons are sensitive to leptin levels despite a lack of leptin receptor expression on GnRH or kisspeptin neurons. It is believed that leptin modulates its reproductive effect by activating nitric oxide—producing neurons in preoptic hypothalamus and ventral premammillary nucleus.

Congenital hypogonadotropic hypogonadism may be inherited as follows:

X-linked, for example, due to anosmin 1, FGF8, SOX10 mutation; passed on to daughters who are carriers of the gene,

autosomal-dominant, for example, PROKR2, PROK2, FEZF1 mutation; vertical transmission to next generation; 50% probability of the offspring to get affected, and

autosomal-recessive, for example, GNRH1, GNRHR, KISS1R, TACR3 mutations; low chances of inheritance to progeny unless a consanguineous relationship.

**Targeted molecular diagnosis and therapy** The gene mutations mentioned later may be associated with a specific phenotypic profile, which can help identify the causal gene for CHH. Table 13.2 shows phenotype—gene mutation correlations, which may be seen in patients with CHH.

**Management of congenital hypogonadotropic hypogonadism** The goal of therapy, for patients with CHH, includes induction of puberty, development of secondary sexual characteristics, and restoration of fertility. The therapeutic options include testosterone (e.g., 200 mg monthly of testosterone enanthate initially and then every 2–3 weeks; lower dose of 50 mg is used in boys aged around 10 years) or estradiol therapy (transdermal estradiol at low dose; 50 µg patches daily followed by 200 mg of progesterone for 14 days): it does not increase gonadal size and has no effect on fertility. Alternatively, pulsatile GnRH therapy, which includes gonadotropins + / – FSH, can be used for induction of puberty and restoration of fertility. hCG (1000–1500 IU twice or thrice a week) combined at a later stage with FSH (75–150 IU twice or thrice a week) have been shown to be efficacious in patients with CHH [16]. Presence of cryptorchidism, small testicular volume, and low inhibin B levels are considered to be poor prognostic factors in the restoration of fertility. Similarly, in girls with CHH, estradiol therapy is started at a low dose (0.05–0.07 µg/kg) with gradual up-titration over 12–18 months. A cyclic progesterone therapy is added when an adequate estrogen level is achieved or, alternatively, when there is menstrual bleeding. The adult dose of estrogen therapy is 50 µg administered transdermally or 1–2 mg oral tablets along with progesterone (200 mg for 2 weeks). Estrogen therapy increases uterine size, which can be monitored radiologically even though it does not have any impact on ovulation or fertility. Gonadotropin and GnRH are usually required for the restoration of ovulatory function and achieving fertility. Around 10%–20% patients with CHH may show a spontaneous reversal of later-mentioned pathophysiological defects with the recovery of HPG axis [17].

### 13.2.3 Genetics of pituitary adenoma

Most of the pituitary adenoma is sporadic although these benign tumors can be part of a familial syndrome. Table 13.3 shows the genetics and clinical characteristics of syndromes associated with pituitary adenoma.

**TABLE 13.3** Genetics and clinical characteristics of syndromes associated with pituitary adenoma.

Syndrome	Gene	Function	Tumor	Extrapituitary manifestations
Familial isolated pituitary adenoma	<i>AIP</i>	Aryl hydrocarbon receptor-interacting protein	GH-secreting tumor	Nil
MEN1 syndrome	<i>MEN1</i>	Menin protein	Prolactinoma, GH, ACTH-secreting tumors NFPA	Primary HPT (manifests by fourth decade) Gastropancreatic NET
MEN4 syndrome	<i>CDKN1B</i>	Cyclin-dependent kinase inhibitor 1B	Same as MEN1	Primary HPT (onset at age >40 years)
McCune–Albright syndrome	<i>GNAS1</i>	Adenylyl cyclase	GH-secreting tumor	Bone dysplasia, precocious puberty, café-au-lait spots
Carney complex	<i>PRKAR1A</i>	Protein kinase A regulatory subunit 1 $\alpha$	GH-secreting tumor	Cushing's syndrome Pigmentation Atrial myxomas

*ACTH*, Adrenocorticotrophic hormone; *GH*, growth hormone; *HPT*, hyperparathyroidism; *MEN*, multiple endocrine neoplasia; *NET*, neuroendocrine tumor; *NFPA*, nonfunctioning pituitary adenoma.

### 13.3 Primary hyperparathyroidism and multiple endocrine neoplasia syndromes

#### 13.3.1 Primary hyperparathyroidism

##### 13.3.1.1 Introduction

Primary hyperparathyroidism (HPT) is one of the commonest endocrine disorders with a prevalence of 1–3:1000 population and a 3:1 female preponderance [18]. It is mostly sporadic although 5%–10% of the patients have familial disease [19]. Primary HPT is characterized by asymptomatic or symptomatic hypercalcemia because of hyperplasia/adenoma of one or more of the parathyroid glands. In developed countries a larger proportion of the patients may be asymptomatic although osteopenia/osteoporosis, renal stones are the most common complications associated with primary HPT. The diagnosis of this condition is based on biochemical evidence of elevated serum calcium and parathyroid hormone (PTH) levels although a normocalcemic variant of HPT may also be encountered in clinical practice [20]. Familial hypocalciuric hypercalcemia (FHH) is a differential diagnosis, which needs to be excluded in patients with primary HPT in view of significant difference in long-term outcomes of these two conditions [21]. Iatrogenic rise in PTH needs to be considered in patients on bisphosphonates, denosumab, and thiazide diuretics.

##### 13.3.1.1.1 Clinical case

A 48-year-old woman was referred to the endocrine clinic with incidentally detected high serum calcium levels while she underwent routine annual blood tests. She was otherwise well. She denied any weight loss, arthralgia, renal colic-related symptoms. She had no previous medical history of any medical illness. She was not taking any medications. Her mother and sister were diagnosed to have high serum calcium levels although no further details were available.

Her general physical and systemic examinations were unremarkable.

Investigations:

Corrected calcium	2.75 mmol/L	(2.2–2.6)
Phosphate	0.65 mol/L	(0.8–1.5)
PTH	12 pmol/L	(1.6–7.2)
Urinary Ca excretion	40 mmol/24 hours	
Calcium:creatinine excretion ratio	0.45	
Ultrasonography of abdomen	Normal	
Bone mineral density scan	T-score 1.1 (average)	

**TABLE 13.4** Clinical characteristics associated with multiple endocrine neoplasia (MEN) syndrome and hyperparathyroidism–jaw tumor (HPT-JT) syndrome.

Diagnosis	Syndrome	Clinical characteristics
Gastrinoma	MEN1 and 4	Recurrent gastritis, peptic ulcer disease, chronic diarrhea
Insulinoma	MEN1 and 4	Episodes of (usually early morning) light headedness, sweating, shakiness; relief of symptoms with glucose
Prolactinoma	MEN1 and 4	Galactorrhea, menstrual irregularity, or secondary amenorrhea
Acromegaly	MEN1 and 4	Arthralgia, sweating, headache, increase in ring or shoe size, dental malocclusion
Pheochromocytoma	MEN2	Paroxysmal or sustained BP, episodic palpitation, sweating, orthostatic giddiness, anxiety
Medullary thyroid carcinoma	MEN2	Neck lump–enlarged lymph nodes
Jaw tumors	HPT-JT	Ossifying mandible or maxilla tumor
Uterine/renal tumors	HPT-JT	Uterine fibroid, renal cysts

BP, Blood pressure.

#### 13.3.1.1.2 Discussion with reflection on the molecular systems underpinning the clinical scenario

Our patient was incidentally detected to have elevated serum calcium as well as PTH levels consistent with a biochemical diagnosis of primary HPT. The signs and symptoms associated with hypercalcemia are generally nonspecific and may include fatigue, malaise, nausea, abdominal pain, and arthralgia (classically described by a famous mnemonic: renal stones, painful bones, abdominal groans, and psychic moans). Majority of the patients diagnosed with primary HPT, especially in developed countries, are relatively asymptomatic. Although primary HPT is mostly sporadic in origin, 5%–10% of patients with primary HPT have a familial disease [19].

To begin with, it is helpful to evaluate the patients diagnosed with primary HPT for a possible associated syndrome complex such as multiple endocrine neoplasia (MEN) syndrome type 1, 2A, and 4 or primary HPT–jaw tumor (HPT-JT) syndrome. Table 13.4 shows clinical characteristics, which may be seen in patients with MEN syndrome and HPT-JT.

Primary HPT remains the presenting clinical manifestation in patients with MEN1 syndrome with almost 100% penetrance by the age of 50 years. These patients may develop primary HPT at a younger age group and may have recurrence of disease. Parathyroid hyperplasia of one or more PTH gland is the classic histopathological feature seen in patients with MEN1 syndrome [22].

In contrast, primary HPT is seen in 20%–30% of patients with MEN2A syndrome. Medullary thyroid cancer (MTC) develops in 50%–60% of patients with MEN2A syndrome and pheochromocytoma is seen in 30%–40% of patients with MEN2A syndrome.

MEN4 syndrome is characterized by a relatively later age (age >45 years) for onset of primary HPT and a milder clinical presentation. It is associated with the mutation of *CDKN1B* gene [22].

HPT-JT syndrome is characterized by development of primary HPT in adolescents and young adults. The patients may develop ossifying tumors of maxilla and mandible. There may be associated renal cysts in 10%–20% patients, and uterine tumors, such as fibroid, have been described in around 40% of women with HPT-JT syndrome. There is an increased risk (15%–20%) of development of parathyroid cancer in patients with HPT-JT syndrome [22]. The classical clinical presentation of parathyroid cancer remains a neck lump and an aggressive parathyroid disease (significant increase in serum calcium and PTH levels).

Table 13.5 shows the genetic and clinical characteristics of syndromes associated with primary HPT.

Nonsyndromic primary HPT may be familial or isolated/ sporadic with both these forms of primary HPT associated with similar genetic mutations (*MEN1*, *CDC73*, and *CaSR* mutations) [23].

Our patient did not have any associated symptoms or signs suggestive of underlying MEN syndrome or HPT-JT syndrome. She had a family history of primary HPT. Therefore she fits into the category of nonsyndromic familial HPT. See Fig. 13.1 for classification of primary HPT based on underlying etiologies.

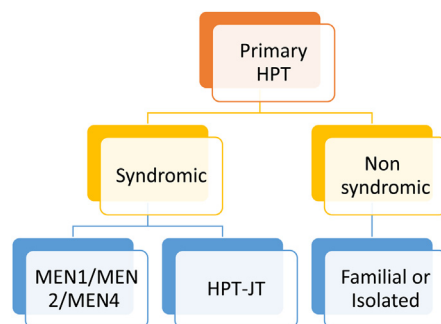
The following clinical characteristics should prompt a clinician to assess for genetic testing in patients with primary HPT:

- Young age at diagnosis (age <50 years)
- Family history of HPT

**TABLE 13.5** Genetic and clinical characteristics of multiple endocrine neoplasia (MEN) syndrome

Syndrome	Gene	1 HPT	Pituitary/pancreas	Pheochromocytoma and MTC
MEN1	<i>MEN1</i> gene (encodes menin)  <i>CDK</i> genes	Present in 100% by the age of 50 years. Mean age of onset: 20–25 years	NFPA  Prolactinoma  Acromegaly NFPNET  Gastrinoma  Insulinoma	Nil
MEN2A	<i>RET</i>  Proto-oncogene (EC cadherin domain)	Present in 20%–30%  Mean age of onset >30 years	Nil	Pheochromocytoma present in 40%–50%  MTC seen in 100%
MEN4	<i>CDKN1B</i> (encodes p27)	Present in 100% although mean age of onset >45 years	Similar to MEN1	Nil

EC, Extracellular; GH, growth hormone; HPT, hyperparathyroidism; IC, intracellular; MTC, medullary thyroid cancer; NFPA, nonfunctioning pituitary adenoma; NFPNET, nonfunctioning pancreatic neuroendocrine tumor; PNET, pancreatic neuroendocrine tumor.

**FIGURE 13.1** Showing classification of 1 HPT based on underlying etiology. HPT, Hyperparathyroidism.

- Recurrent disease
- Multiglandular involvement
- History of recurrent gastritis, malabsorption, gastric peptic ulcer disease
- History of paroxysmal or sustained rise in BP and/or thyroid cancer
- Jaw tumors, renal or uterine tumor onset at a young age

In our patient the young age of presentation (age <50 years) and strong family history of hypercalcemia are the factors, which weigh in favor of further genetic analysis. She had no significant past medical history and was not on any medication (such as lithium, bendroflumethiazide, and denosumab), which could influence calcium and PTH results. The gene analysis panel, which needs to be carried out in her case, should include *MEN1*, *CDK*, and *cell division cycle 73* (*CDC 73*).

An important differential diagnosis, which needs to be considered, includes FHH even though a calcium:creatinine excretion of >0.2 makes this diagnosis unlikely. Inactivating mutation of calcium-sensing receptor (*CaSR*) gene mutation is associated with development of FHH subtype 1, which is characterized by hypercalcemia and hypocalciuria. It has an autosomal dominant inheritance pattern. The PTH levels may be high or normal. The hallmark of this condition remains a urinary calcium:creatinine (UCaCr) clearance ratio of <0.01, although the following factors may influence UCaCr clearance ratio [24]:

- Ethnic origin (lower UCaCr clearance ratio is seen in individuals of Africo-Caribbean origin)
- Medications (loop diuretic and lithium)
- eGFR
- 25-Hydroxy vitamin D levels



**TABLE 13.6** Genetics and prevalence of familial hypocalciuric hypercalcemia (FHH) subtypes.

FHH subtype	Gene	Prevalence (%)
Type 1	Loss of function mutation in <i>CaSR</i>	65
Type 2	<i>AP2S1</i>	>10
Type 3	<i>GNA11</i>	>20

*CaSR*, Calcium-sensing receptor.

Moreover, there may be an overlap in UCaCr clearance ratio (0.01–0.02) in individuals with FHH and primary HPT [24]. FHH subtype 2 is characterized by mutation of *AP2S1* gene, which encodes the adaptor protein-2 sigma subunit; FHH subtype 3 is characterized by loss of function mutation in *GNA11*, which encodes for G-protein subunit  $\alpha$ -11. These patients have a low phosphate level in addition to lower calcium levels. Patients with FHH subtype 3 are more likely to be symptomatic as compared to FHH1. There may be associated features such as short stature, learning defects, and pancreatitis [23]. Table 13.6 shows the genetics and prevalence of the 3 subtypes of FHH.

#### 13.3.1.1.3 Overview of the relevant molecular systems underpinning the clinical scenario

*MEN1* gene encodes for protein menin, which is involved in the suppression of tumorigenesis. Mutation in *MEN1* gene results in an altered menin protein resulting in increased cellular proliferation and cell division. This results in hyperplasia and/or adenoma formation involving pituitary, parathyroid, and pancreas. In <2% patients with MEN1 syndrome, there may be an unknown, detectable *MEN1* gene mutation. Instead, these patients may harbor a mutation in *CDKI* (cyclin-dependent kinase inhibitors) gene family. The *CDK* family is involved in the regulation of cell-cycle progression and cell growth. The *CDK* gene family includes *CDKN1A* encoding for p21, *CDKN1B* encoding for p27, *CDKN2B* encoding for p15, and *CDKN2C* encoding for p18. A loss of function mutations in *CDK1* family genes promotes tumorigenesis.

Mutations in extracellular cadherin domain and intracellular tyrosine kinase domain of *RET* (rearranged during transfection) proto-oncogene are associated with the development of MEN2A and MEN2B syndrome, respectively.

MEN4 syndrome is caused by inactivating mutations in the *CDKN1B* gene. *CDKN1B* gene encodes p27, which is a key regulator of cell-cycle progression. Disruption of p27 results in uncontrolled cell-cycle progression in neuroendocrine cells.

Cyclin proteins play a key role in the regulation and progression of cells through the cell cycle. These activate cyclin-dependent kinases (CDKs). *CDC73* gene encodes for protein parafibromin, which is predominantly present within the nucleus of cells. Parafibromin protein is involved in gene transcription and acts as a tumor-suppressor gene. Extranuclear parafibromin is involved in the organization of cellular cytoskeleton. *CDC73* mutations are associated with the development of familial HPT, HPT-JT syndrome, and parathyroid cancer [23,25,26]. Table 13.7 shows the spectrum of disorders associated with *CDC73* mutations.

Genetic screening in the select group of individuals is helpful in confirming the diagnosis and help screen for associated tumors (e.g., MEN1, MEN2A, and MEN4 or HPT-JT syndrome). It can also help in identifying the appropriate surgical approach for the patients, for example, for MEN syndrome in patients with multiglandular involvement, open neck exploration is preferred over minimally invasive elective parathyroidectomy. Genetic screening of the relatives, who may be harboring the mutation, is useful for early detection and prompt therapeutic intervention.

#### 13.3.1.2 Management of patients with primary hyperparathyroidism

The management of patients with primary HPT can be broadly divided into:

##### 13.3.1.2.1 Surgical management

According to the National Institute of Health's (NIH) consensus criteria [27,28], elective parathyroidectomy remains the preferred management option, which can be seen in Table 13.8.

##### 13.3.1.2.2 Medical management

Asymptomatic patients not fulfilling NIH consensus criteria are generally managed conservatively [27,28]. The key focus should be on the following:

- To replace 25(OH) vitamin D if patients have a coexistent deficiency
- Bisphosphonates for osteoporosis

**TABLE 13.7** Spectrum of disorders associated with CDC73 mutations.

Condition	Onset	Clinical characteristics	Carrier surveillance
FIHPT	Age <45 years	More severe hypercalcemia as compared to sporadic HPT	Annual Ca/PTH/25(OH) vitamin D
HPT-JT syndrome	Late adolescence	Ossifying fibromas in the jaw Renal cysts Uterine tumors	Panoramic dental X-rays every 5 years (from age 10) Renal ultrasound every 5 years
Parathyroid cancer	Age <50 years	Palpable neck lump Severe hypercalcemia	Annually

Ca, Serum calcium; FIHPT, familial isolated hyperparathyroidism; HPT-JT, hyperparathyroidism–jaw tumor syndrome; PTH, parathyroid hormone.

**TABLE 13.8** Guidelines for parathyroid surgery in asymptomatic primary hyperparathyroidism (PHPT) (JCEM, Oct 2014).

Total corrected serum calcium	2.85 mmol/L (> 3 mmol/L-UK)
Creatinine clearance	<60 mL/min (no UK guidelines)
Age	<50 years
24-h urinary calcium	10 mmol/day (> 400 mg/day) and increased risk of stone formation Presence of nephrolithiasis or nephrocalcinosis
Bone mineral density	T-Score < −2.5 at lumbar spine, total hip, femoral neck, or distal one-third radius (Z-score in men <50 years and premenopausal women)
Medical surveillance	Vertebral or fragility fracture Not possible/not desirable

**TABLE 13.9** Guidelines for monitoring patients managed conservatively/nonoperatively in asymptomatic primary hyperparathyroidism (PHPT).

Measurement	United Kingdom	NIH (United States)
Serum creatinine	Every 6 months	Every 12 months (eGFR)
Abdominal X-ray/USS	Every 3 years	—
Bone density	Every 2–3 years (three sites)	Every 1–2 years
Serum calcium	Every 6 months	Every 12 months
PTH/urinary calcium	Every 12 months	—

PTH, Parathyroid hormone; USS, Ultrasound scan.

- Cinacalcet is a calcimimetic agent, which lowers serum calcium even though its impact on PTH levels is modest, if any. It has been approved by NICE (National Institute of Clinical Excellence) guidelines to be used in symptomatic hypercalcemia (serum calcium >2.85 mmol/mol) patients who are deemed unfit for surgery. It has not been shown to be associated with any significant impact on bone markers of osteoporosis. The limitation, in its use, remains high cost and associated gastrointestinal (GI) side effects. There is no long-term data showing any significant impact on reduction in end-organ damage (nephrolithiasis and osteoporosis) due to primary HPT by Cinacalcet use.

Table 13.9 shows the guidelines for monitoring of patients with asymptotic primary HPT who are managed conservatively (nonsurgically).

## 13.4 Chronic hypocalcemia and hypophosphatemia

### 13.4.1 Chronic hypocalcemia

#### 13.4.1.1 Introduction

PTH and 1,25(OH)<sub>2</sub> D play a key role in maintaining serum calcium levels. The following conditions can result in hypocalcemia:

- Hypoparathyroidism or resistance to PTH action
- 1,25(OH)<sub>2</sub> D deficiency or resistance to 1,25(OH)<sub>2</sub> D action
- Acute deposition of calcium as seen after rhabdomyolysis, tumor lysis syndrome, acute pancreatitis

Signs and symptoms related to hypocalcemia depend upon the severity and duration of hypocalcemia. These include carpopedal spasms (Trousseau's sign) or peri-oral paresthesia (Chvostek's sign). In addition, acute hypocalcemia can lead to generalized seizures. Hypocalcemia is also associated with prolongation of QT interval, sub-capsular cataract formation, basal ganglion calcification, and dry and flaky appearance of skin [29,30]. A systematic approach in the clinical evaluation of patients presenting with hypocalcemia helps in identifying the underlying etiology and alleviating long-term morbidity.

#### 13.4.1.2 Clinical case

A 26-year-old woman was referred from dental clinic in view of low serum calcium levels. She was under follow-up of the dental clinic for her loss of teeth since the age of three and poor periodontal status. She had past medical history of epilepsy. There was no history of surgical procedure involving neck or exposure to radiation. There was no family history of hypocalcemia. She was not on any medication.

On examination, she had short stature, round face, short neck, and a short fourth finger.

Investigations:

Corrected serum calcium	1.92 mmol/L	(2.2–2.6)
Serum phosphate	1.7 mol/L	(0.8–1.5)
Serum PTH	8.8 pmol/L	(1.6–7.2)
Serum vitamin D	72 nmol/L	
TSH	6.8 mU/L	(0.4–4.5)
FT4	22.4 pmol/L	(10–20.5)
eGFR	>90 mL/min	

#### 13.4.1.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

This young woman has hypocalcemia and hyperphosphatemia although PTH remains inappropriately high. Her biochemical investigation results are consistent with possible PTH resistance. She has a background history of poor dentition as well as seizures, which is suggestive of chronicity of hypocalcemia. She also has classic phenotypic features suggestive of pseudohypoparathyroidism (PHP). PHP has two main subtypes:

- Type 1A characterized by phenotypic features, such as short stature, round face, short fourth and fifth metacarpal, and metatarsal bones, along with biochemical features of low calcium, raised phosphate, and elevated PTH levels. It is also known as Albright's hereditary osteodystrophy.
- Type 1B, which is characterized by biochemical abnormalities, such as hypocalcemia, hyperphosphatemia, and raised PTH levels although phenotypic features seen in Type 1A subtype are absent.

It is worth remembering that PHP may be associated with other endocrine abnormalities such as primary hypothyroidism and/or abnormalities of reproductive function.

PHP type 1 is caused by loss of function mutation of the gene encoding stimulatory G protein  $\alpha$ -subunit (GNAS). This results in resistance to the action of PTH on binding with G protein  $\alpha$ -subunit. As the G protein-coupled receptors (GPCR) mediate action of other hormones, such as TSH, gonadotropins, and GH,

**TABLE 13.10** Genetic and clinical characteristics associated with *GNAS* mutations.

Characteristics	PHP 1A	PHP 1B	PPHP
Low Ca and high PTH	Present	Present	Absent
Phenotypic features including short fourth and fifth metacarpal	Present	Absent	Present
Genetic mutation	<i>GNAS</i> coding region	<i>GNAS</i> regulatory region	<i>GNAS</i> coding region (paternally transmitted)
Generalized hormone resistance	Present	Absent	Absent

*PHP*, Pseudohypoparathyroidism; *PPHP*, pseudopseudohypoparathyroidism; *PTH*, parathyroid hormone.

patients with *GNAS* mutation may present with clinical and biochemical features of hormone resistance to more than one hormone [31]. In our patient, FT4 levels are elevated with raised TSH consistent with possible TSH resistance.

It is inherited as an autosomal-dominant trait. There is genomic imprinting of the *GNAS* locus, which results in PTH (and other GPCR-coupled hormones) resistance if the mutant allele is derived maternally. If the mutant allele is acquired from the father, it leads to classical phenotypic features associated with PHP type 1 (short stature, round face, shortened fourth and fifth metatarsals) even though there is an absence of hormone resistance. This condition is also known as pseudopseudohypoparathyroidism [31,32].

PHP type 1B patients are characterized by loss of parent-specific *GNAS* regulatory region mutation or deletion. The *GNAS* coding region in these patients is normal. Majority of patients with PHP 1B are sporadic. Table 13.10 shows the genetic and clinical characteristics associated with *GNAS* mutations:

#### 13.4.1.2.2 Management of patients with pseudohyperparathyroidism

The main differential diagnosis of PHP remains secondary hyperparathyroidism, which is commonly due to renal impairment or vitamin D deficiency. The phenotypic features, such as short stature, round face, and shortened fourth and fifth metatarsal, are supportive of the diagnosis of PHP type 1A. In addition, hyperphosphatemia is seen in PHP in contrast to hypophosphatemia as commonly observed in vitamin D deficiency. Ellsworth–Howard test can be used to confirm the diagnosis of PHP. It involves an infusion of synthetic human PTH intravenously and serial measurement of cAMP. A failure to increase urinary cAMP (<300 nmol/L) after PTH infusion is supportive of PTH resistance. Management of these patients includes oral calcium supplements and active metabolites of vitamin D such as calcitriol.

### 13.4.2 Chronic hypophosphatemia

#### 13.4.2.1 Introduction

Chronic hypophosphatemia is associated with the development of bony deformities including rickets and osteomalacia. The etiologies leading to hypophosphatemia include poor nutrition, malabsorption, iatrogenic factors (e.g., diuretics, steroids, antacids), and renal tubular disorders although rarely does chronic hypophosphatemia have a genetic basis (X-linked and autosomal-dominant hypophosphatemia).

#### 13.4.2.2 Clinical case

A 16-year-old girl was reviewed in a pediatric endocrinology transition clinic with weakness and bone pain. She had a background history of delay in developmental milestones and bone deformities. There was no history suggestive of malabsorption or eating disorder. She was not on any medication. Her mother and grandmother were also known to have bone deformities.

On examination, her BP was 110/76 mm of Hg. She had short stature and bone deformities. There was no evidence of proximal myopathy or easy bruisability.

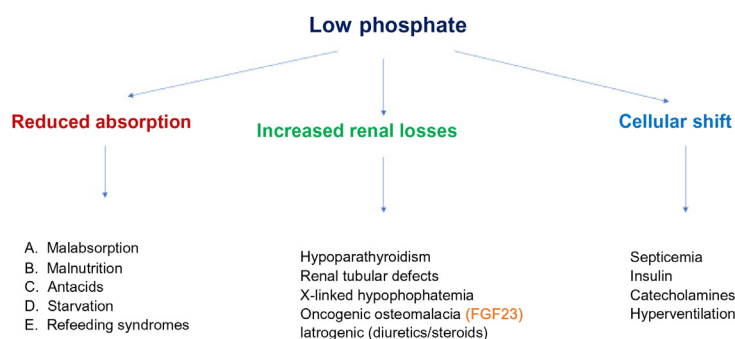
## Investigations:

Corrected serum calcium	2.3 mmol/L	(2.2–2.6)
Serum phosphate	0.35 mol/L	(0.8–1.5)
Serum PTH	8.8 pmol/L	(1.6–7.2)
25(OH) vitamin D	65 nmol/L	
eGFR	>90 mL/min	

## 13.4.2.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

This young girl has hypophosphatemia and elevated PTH levels in the presence of normal calcium, vitamin D levels as well as normal renal function. The bone pain, short stature, and deformities point toward a defect in bone mineralization (e.g., rickets or osteomalacia).

The following algorithm is useful while evaluating patients with chronic hypophosphatemia:



Our patient does not have any iatrogenic factor that could have contributed to her low phosphate levels. In the absence of a history of malnutrition, malabsorption, eating disorders, and any obvious renal defect, our patient likely has chronic hypophosphatemia, which likely has familial/genetic origin based on the family's history of skeletal deformities.

The differential diagnosis for patients presenting with hypophosphatemia and bone mineralization defect is shown in Table 13.11.

Based on the clinical and biochemical profile, the most likely diagnosis for our patient is X-linked hypophosphatemia (XLH) and there is a need for further genetic testing to confirm this diagnosis. Autosomal-dominant, autosomal-recessive as well as renal transporter gene (*SLC34A1*, *SLC34A3*, *SLC20A2*) mutation-related forms of hypophosphatemic rickets are also known although these remain extremely rare [33].

## 13.4.2.2.2 Overview of the relevant molecular systems underpinning the clinical scenario

XLH is a rare genetic disorder characterized by hypophosphatemia, reduced 1,25(OH)<sub>2</sub> levels and bone mineralization defects. The clinical features are variable with the disease features being more severe in males. XLH is caused by mutation in *PHEX* gene, which is involved in renal phosphate transport. *PHEX* gene is hypothesized in playing an indirect role in inactivation of FGF23. *PHEX* gene mutation leads to an increased level of circulating FGF23 resulting in renal phosphate losses [34]. Patients are more prone to develop nephrocalcinosis apart from developing male deformities associated with rickets and osteomalacia.

## 13.4.2.2.3 Management of patients with X-linked hypophosphatemia

The management of these patients involves a combined therapy including phosphate (1–4 µg/day) along with calcitriol (1–3 µg/day) [35]. Burosumab, which is a monoclonal antibody to FGF23, has been approved by FDA (Federal Drug Agency) for use in patients with XLH although it has not been approved by NICE for use in the United Kingdom.

**TABLE 13.11** Genetic, biochemical, and clinical characteristics of various disorders associated with hypophosphatemia and osteomalacia.

<b>Osteomalacia and hypophosphatemia</b>							
<b>Condition</b>	<b>Defect</b>	<b>Ca</b>	<b>PO<sub>4</sub></b>	<b>25(OH) D</b>	<b>1,25(OH)<sub>2</sub> D</b>	<b>PTH</b>	<b>Clinical features</b>
Vitamin D deficiency	Low vitamin D	Low	Low	Low	Low	High	P. myopathy, falls, bone pain
VDDR type 1	1- $\alpha$ -hydroxylase deficiency (AR)	Low	Low	N	Low	High	Rickets, myopathy, enamel hypoplasia Rx: alfacalcidol/calcitriol
VDDR type 2	Vitamin D–receptor mutation (AR)	Low	Low	N	High	High	Osteomalacia, may have alopecia. Rx: high-dose calcitriol
X-linked hypophosphatemia (vitamin D–resistant rickets)	<i>PHEX</i> mutation	N	Low	N	N	N/High	Rickets and osteomalacia Short stature No myopathy Rx: PO <sub>4</sub> + alfacalcidol or calcitriol (to suppress PTH)

AR, Autosomal-recessive; Ca, serum calcium; PO<sub>4</sub>, serum phosphate; PTH, parathyroid hormone; Rx, treatment; VDDR, vitamin D deficiency rickets.

## 13.5 Primary hyperaldosteronism

### 13.5.1 Introduction

Primary hyperaldosteronism (PA) remains an underdiagnosed endocrinological condition despite it being believed to be the underlying diagnosis in 5%–12% of patients with hypertension based on results from epidemiological and observational studies [36–38]. The prevalence of PA is even higher in patients with resistant hypertension with 20% of such patients having PA based on results of study by Douma et al. in 2008 [39]. The diagnosis of PA remains challenging as biochemical screening for this condition relies on measurement of serum aldosterone-to-renin-activity-ratio (ARR) although this biochemical investigation is not freely available and a vast majority of commonly used antihypertensive medications having an impact on measurement of ARR. PA can be broadly classified into the following main categories:

- Aldosterone-producing adenoma (APA)
- Bilateral idiopathic hyperplasia (IHA)
- Familial hyperaldosteronism (FH)

APA and IHA account for 35% and 60% of all cases of PA although it may be challenging to distinguish between the two based on clinical characteristics. FH accounts for <5% of the cases of PA. It is further subdivided into the following:

- FHA type 1 or glucocorticoid-remediable aldosteronism
- FH type 2
- FH type 3

PA is associated with a higher risk of development of end-organ damage associated with hypertension as compared to patients with essential hypertension. In addition, these patients are at a higher risk of developing cardiovascular events such as myocardial infarction, atrial fibrillation, and stroke. Early recognition and therapeutic intervention is important in patients with PA to prevent long-term morbidity and mortality associated with untreated disease.

### 13.5.2 Clinical case

A 33-year-old man was referred to an endocrine clinic in view of uncontrolled BP despite being on three anti-hypertensive medications (Ramipril, Amlodipine, and Doxazosin). He was diagnosed to have high BP about 18 months ago. He was a nonsmoker and an occasional social drinker. He had no family history of high BP or hemorrhagic stroke.



On examination, his body mass index was 25.5 kg/m<sup>2</sup>. His pulse rate of 84/minute and BP was 160/90 mm of Hg. He did not have papilledema. The systemic examination was normal. There was no audible renal bruit.

Investigations:

Na	142 mmol/L	(135–145)
K	3.6 mmol/L	(3.5–5.5)
Creatinine	80 µmol/L	(60–115)
Serum aldosterone	1880 pmol/L	(100–800)
Plasma renin activity	0.1 pmol/L	(0.5–3.5)
CT abdomen	2.5 cm Left-sided adrenal adenoma	

### 13.5.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

This man has been diagnosed to have high BP at a young age and it remains uncontrolled despite him being on three different antihypertensive medications. Renal and endocrine diseases remain the commonest etiologies leading to secondary hypertension; although in around 85% of patients with hypertension, no underlying cause could be detected. PA remains one of the commonest endocrine etiologies leading to secondary hypertension. Patients with hypertension need to be screened for PA in presence of one or more of the following clinical characteristics:

- Young age at the time of diagnosis (<40 years)
- Refractory hypertension
- Severe hypertension (> 160/100 mm of Hg)
- Presence of hypokalemia and/or metabolic alkalosis
- Strong family history of hypertension and/or hemorrhagic stroke
- Incidental adrenal lesion in a patient with hypertension

The young age and uncontrolled hypertension in our patient are strong-enough reasons to screen for PA. It is worth remembering that hypokalemia is present only in approximately 35% of patients with PA. The screening test for PA includes measurement of serum ARR. The following medications may have an impact on ARR:

- β-Blockers and aldosterone antagonists (spironolactone, eplerenone) can result in false positive results and need to be stopped for 2 and 6 weeks, respectively.
- Angiotensin-converting enzyme inhibitors and calcium channel blockers can result in false negative results although stopping these medications is not recommended while screening for PA.

In our patient, initial screening test (ARR) suggests an elevated aldosterone and suppressed aldosterone consistent with a possible underlying diagnosis of PA. An oral or intravenous saline-loading test with repeat measurement of ARR can be used to confirm the biochemical findings of the initial screening test although fludrocortisone–dexamethasone suppression test is considered a gold standard for the diagnosis of PA by some of the experts. Our patient has a 2.5 cm left-sided adrenal adenoma, which potentially is the source of excessive aldosterone secretion. In absence of a family history of hypertension and/or hemorrhagic stroke, it is most likely an isolated PA secondary to an aldosterone-secreting adrenal adenoma. At times, it is clinically challenging to distinguish between an aldosterone-secreting adenoma and bilateral adrenal hyperplasia. In such patients, selective adrenal venous sampling can help differentiate between the aforementioned conditions although it remains an invasive procedure, which may only be available in tertiary-care centers.

### 13.5.2.2 Overview of the relevant molecular systems underpinning the clinical scenario

In the last decade, somatic mutations have been diagnosed in around 50% patients with adrenal adenoma [40,41]. Table 13.12 shows the genetics and clinical characteristics of the subtypes of APA based on underlying somatic mutation.

FH is further subdivided into three categories. Table 13.13 shows the genetics and clinical characteristics of subtypes of FH.

For a 33-year-old man with uncontrolled hypertension, the most likely diagnosis is APA in case hypertension or hemorrhagic stroke is absent in the family history. The most likely somatic mutations in such patients are *KCNJ5*, *ATP1A1*, *ATP2B3*, and *CACNA1D* as shown in Table 13.12.

**TABLE 13.12** Genetics and clinical characteristics of subtypes of aldosterone-producing adenoma.

Gene	Encodes	Prevalence	Clinical characteristics
<i>KCNJ5</i>	Kir 3.4 potassium channel	40%	Female preponderance Younger patients
<i>ATP1A1</i>	Sodium/potassium ATPase		Male preponderance Severe disease Relatively smaller adenomas
<i>ATP2B3</i>	Calcium ATPase		
<i>CACNAID</i>	Voltage-dependent C-type calcium channel		

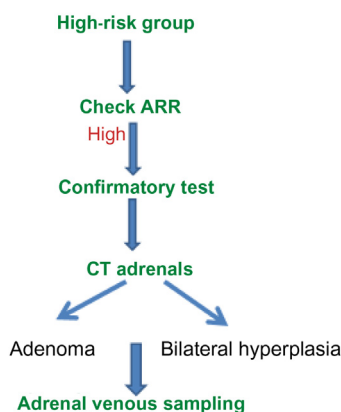
**TABLE 13.13** Genetics and clinical characteristics of subtypes of familial hyperaldosteronism (FH).

Subtype	Gene	Molecular mechanism	Clinical characteristics
FH1	<i>CYP11B1</i> and <i>CYP11B2</i>	Chimeric gene duplication due to crossing-over between promoter sequence of <i>CYP11B1</i> , which encodes for 11 $\beta$ -hydroxylase, and coding sequence of <i>CYP11B2</i> , which encodes aldosterone synthase	<ul style="list-style-type: none"> <li>Mineralocorticoid production under influence of ACTH</li> <li>Hemorrhagic stroke at a young age</li> <li>Increased 18-hydroxycortisol and 18-oxocortisol</li> <li>Responds to glucocorticoids</li> </ul>
FH2	Unclear	Based on linkage studies, an association has been shown with chromosome region 7p22	Autosomal-dominant condition
FH3	<i>KCNJ5</i>	Germline mutation in <i>KCNJ5</i> leading to increased sodium conductance and cell depolarization, which triggers calcium entry and aldosterone production	<ul style="list-style-type: none"> <li>Refractory hypertension</li> <li>Patients may need adrenalectomy</li> </ul>

### 13.5.3 Management of patients with primary hyperaldosteronism

The patients with PA due to an underlying APA should be considered for elective unilateral laparoscopic adrenalectomy. The patients with APA, in whom surgery is deemed inappropriate or not feasible, should be treated medically with aldosterone antagonists (spironolactone or eplerenone). A similar pharmacotherapeutic approach is adopted for patients with PA due to IHA and FH2. The patients with FH1 usually respond to glucocorticoid therapy (dexamethasone 0.25–0.5 mg once a day orally) although aldosterone antagonists are considered an alternative in such patients. Patients with FH3 usually have refractory hypertension, which may be unresponsive to aldosterone antagonists. These patients may need elective adrenalectomy. The following algorithm shows the management approach for patients with PA.

#### Approach to diagnosis



## 13.6 Congenital adrenal hyperplasia, apparent mineralocorticoid excess, and renal tubular defects

### 13.6.1 Congenital adrenal hyperplasia

#### 13.6.1.1 Introduction

Congenital adrenal hyperplasia (CAH) is a group of inherited disorders due to deficiency/absence of enzymes involved in cortisol production in zona glomerulosa of adrenal glands. The commonest enzymatic defect leading to CAH involves absence of 21- $\alpha$  hydroxylase, which catalyzes conversion of 17-OH progesterone (17-OHP) to 11-deoxycortisol. Patients with absence of 21- $\alpha$  hydroxylase (also known as classic CAH) may present in neonatal period with poor oral intake, failure to thrive, and low BP. These neonates are hyponatremic and hypovolemic on presentation. The diagnosis can be confirmed by measuring 17-OHP levels, which rise. Classic CAH has an autosomal-recessive inheritance with 1 in 60 people carrying the gene. The management includes steroid therapy. A typical regimen includes prednisolone (5–7.5 mg in one or two divided doses). Despite adequate steroid replacement, fertility in patients with classic CAH is reduced by 20%–30%. The follow-up should include titration of the steroid dose to keep 17-OHP about twice as normal and plasma renin levels in midrange [42,43]. Patients on long-term steroid therapy should be evaluated for signs and symptoms of iatrogenic excess-steroid intake. These patients should be monitored for possible osteoporosis by regular bone-density measurement. In addition, around 25% of the female patients with classic CAH may have gender identification issues. Psychosexual counseling is helpful in such scenarios. There are 50% chances of the transmission of gene mutation to the offspring. It is advisable to treat a mother with CAH with dexamethasone as it crosses placenta and can suppress fetal androgens if raised although if it is a male or a genetically unaffected fetus, steroid therapy can be stopped.

Patients with partial deficiency of 17-OHP (also known as nonclassic CAH) generally present in late adolescence. Females with nonclassic CAH have menstrual irregularities, acne, and biochemical evidence of hyperandrogenism. Nonclassic CAH remains an important differential diagnosis of polycystic ovarian syndrome (PCOS), although elevated 17-OHP levels help distinguish it from the former. A few rarer subtypes of CAH include the following:

- 17-Hydroxylase deficiency
- 11 $\beta$ -Hydroxylase deficiency

#### 13.6.2 Clinical case

A 17-year-old girl was referred to an endocrinology clinic with primary amenorrhea. She had no history of headache, visual disturbance, or galactorrhea. She had no developmental delay and had attained puberty at an age of 12 years. Her body weight had been stable, and she denied any nausea, vomiting, or alteration in bowel habits. She had no past history of any medical illness. She was not on any medications. There was no family history of note.

On examination, her BP was 150/96 mm of Hg and her pulse rate was 78/minute. On examination, her body mass index was 23.5 kg/m<sup>2</sup>. She had excessive facial and body hair, although there was no evidence of proximal myopathy, easy bruisability, or abdominal stria.

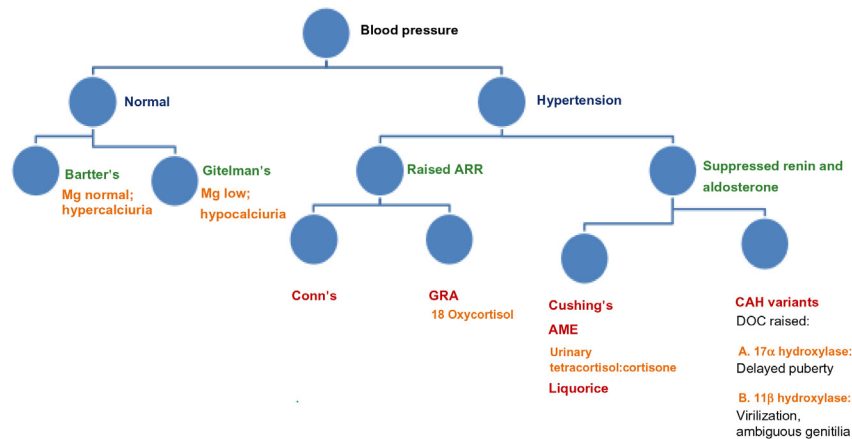
Investigations:

Serum Na	142 mmol/L	(135–145)
Serum K	3.2 mmol/L	(3.5–5.5)
Serum creatinine	80 $\mu$ mol/L	(60–115)
Serum aldosterone	74 pmol/L	(100–800)
Plasma renin activity	0.1 pmol/L	(0.5–3.5)
Serum testosterone	3.5 nmol/L	(<1.5)
Serum LH	8.5 U/L	(3–8)
Serum prolactin	220 mU/L	(<350)
TSH	3.8 mU/L	(0.4–4.5)
FT4	15.4 pmol/L	(10–20.5)
Estradiol	120 pmol/L	(17–260 preludeal)

### 13.6.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

Our patient has primary amenorrhea and clinical features of hyperandrogenism. She is hypertensive and has hypokalemia. Based on the history and examination, there is no obvious GI cause for her hypokalemia. The differential diagnosis for patients presenting with hypokalemia due to non-GI causes is as shown in the algorithm:

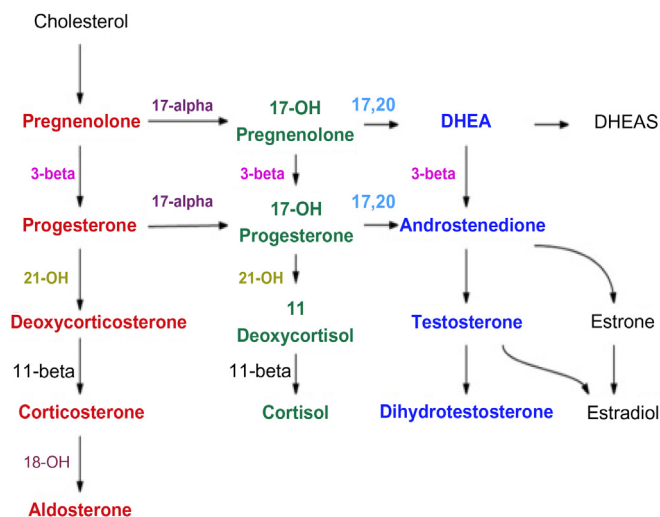
#### Algorithm for approach: low K (Non-GI causes)



The presence of hypertension narrows down the possibilities of hypokalemia to mainly endocrine causes and a suppressed aldosterone; plasma renin activity is suggestive of a possible atypical variant of CAH including 17-hydroxylase deficiency, which is characterized by delayed puberty, absence of secondary sexual characteristics, or primary amenorrhea. These patients are cortisol-deficient, yet do not experience adrenal crises [44,45]. The precursor hormones, such as corticosterone, are elevated.

### 13.6.2.2 Management

The management includes low-dose steroid therapy (e.g., dexamethasone 0.25–1 mg/day). The following is a schematic depiction of adrenal steroid hormone synthesis:



### 13.6.3 Apparent mineralocorticoid excess and renal tubular defects

#### 13.6.3.1 Introduction

Apparent mineralocorticoid excess and renal tubular defects remain important differential diagnosis for patients presenting with hypokalemia and/or hypertension. Apparent mineralocorticoid excess may be of genetic (autosomal-recessive) or iatrogenic (intake of licorice) in origin. Mineralocorticoid receptors present in renal tubules have equal affinity for cortisol and aldosterone. Circulating cortisol levels are 100 times as compared to aldosterone levels. Inactivating mutation of 11 $\beta$ -hydroxysteroid dehydrogenase (11 $\beta$ -HSD) gene, which encodes for type 2 enzyme, converts cortisol to cortisone and protects the MR from activation by circulating cortisol. Enzymatic defects in 11 $\beta$ -HSD can lead to apparent mineralocorticoid excess characterized by early onset of hypertension and hypokalemia. These patients have excess urinary tetrahydrocortisol-to-tetrahydrocortisone ratio [46]. Table 13.14 shows the clinical characteristics and management of conditions leading to apparent mineralocorticoid excess.

#### 13.6.3.2 Clinical case

A 28-year-old woman was referred to an endocrinology clinic due to incidentally detected low serum-potassium levels. She denied any history of nausea, vomiting, abdominal pain, alteration of bowel habits, or weight loss. She had no significant past medical history of any illness. She was not on any regular medications.

On examination, her body mass index was 24 kg/m<sup>2</sup>. Her BP was 130/76 mm of Hg. The general physical and systemic examinations were unremarkable.

##### Investigations

Serum Na	142 mmol/L	(135–145)
Serum K	3.2 mmol/L	(3.5–5.5)
Serum creatinine	80 $\mu$ mol/L	(60–115)
Serum aldosterone	1250 pmol/L	(100–800)
Plasma renin activity	8.5 pmol/L	(0.5–3.5)
Deoxycorticosterone	Normal	
Urinary calcium excretion	High	

#### 13.6.3.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

This young woman was incidentally detected to have hypokalemia. There were no clinical features suggestive of gastrointestinal etiology, for hypokalemia, in her case. On examination, her BP was normal, which was more in keeping with a likely nonendocrine etiology of hypokalemia. Renal tubular transport defects, such as Bartter's

**TABLE 13.14** Clinical characteristics and management of conditions leading to mineralocorticoid excess.

Mineralocorticoid-excess syndrome		Mechanism	Investigation	Management
Condition	Clinical features			
Apparent mineralocorticoid excess	Failure to thrive, polydipsia, HTN (type 1) or its milder form (type 2)	Deficiency of 11 $\beta$ -HSD2; inactivating mutation	Urinary tetrahydrocortisol: tetrahydrocortisone ratio increased	Dexamethasone
Licorice ingestion	History of intake of sweets, cough syrup, herbals	Glycyrrhetic acid inhibits 11 $\beta$ -HSD2	As per clinical scenario	Stop the offending agent
CAH variant (17 $\alpha$ -hydroxylase)	Delayed puberty (women)	17 $\alpha$ -Hydroxylase-deficient	Deoxycortisone increased	Steroids
CAH variant (11 $\beta$ -hydroxylase)	Virilization; ambiguous genitalia	11 $\beta$ -Hydroxylase-deficient	Deoxycortisone increased	Steroids
Cushing's disease	Proximal myopathy, stria, bruising	MR activity of cortisol	ODST/UFC/Salivary cortisol	Surgical

CAH, Congenital adrenal hyperplasia; HTN, hypertension; ODST, overnight dexamethasone suppression test; UFC, 24-h urinary free cortisol.

**TABLE 13.15** Pathogenesis and clinical characteristics of renal tubular transport defects associated with hypokalemia.

<b>Pseudoaldosteronism (renal tubular transport defects)</b>				
<b>Condition</b>	<b>Clinical features</b>	<b>Biochemistry</b>	<b>Mechanism</b>	<b>Management</b>
Bartter's syndrome	BP: normal	Low K	Na-K-2Cl cotransporter (ascending loop of Henle) loss of function	K replacement
	Early onset	Salt wasting		PG synthesis inhibitors (indomethacin)
	Seizures	Hypercalciuria		
	Tetany	Raised renin and aldosterone		
Gitelman's syndrome	BP: normal	Low Mg, K, and Ca leading to hypocalciuria	Na-Cl transporter (DCT) loss of function <i>SCL12A3</i> gene	K and Mg replacement
	Older age of presentation	Raised renin and aldosterone		K-sparing diuretics
Liddle's syndrome	Hypertensive	Low K/suppressed renin and aldosterone	Constitutive activation of Na channel (distal nephron)	Amiloride (not spironolactone)

BP, Blood pressure; Ca, serum calcium; DCT, distal convoluted tubule; K, potassium; Mg, magnesium; PG, prostaglandin.

and Gitelman's syndrome, should be considered as differential diagnoses in patients presenting with hypokalemia and normal BP. The presence of hypercalcemia in our patient was consistent with a diagnosis of Bartter's syndrome although it is a diagnosis of exclusion and genetic tests are useful to confirm it. Table 13.15 shows the characteristics of renal tubular transport defect associated with hypokalemia [47]:

#### 13.6.3.2.2 Management of apparent mineralocorticoid excess and renal tubular defects

The management of these conditions depends upon precise underlying etiology and is summed up in Tables 13.14 and 13.15.

## 13.7 Pheochromocytoma and paraganglioma

### 13.7.1 Introduction

Pheochromocytomas (PHEOs) and nonhead and neck paragangliomas (non-HN PGLs) are tumors originating from the sympathetic nervous system. PHEOs arise from adrenal medulla while NHL PGL arises from ganglion of sympathetic nervous system. In contrast, head and neck (HN) PGL originates from the parasympathetic nervous system and is typically localized above the mediastinum. PHEOs and PGL are relatively rare tumors, with a prevalence of 2–5/million population, although majority of these tumors are believed to remain underdiagnosed. The prevalence of PHEOs and PGL, based on autopsy series, is believed to be 250–1300 cases per million [48,49]. PHEOs and non-HN PGL may present with clinical features of catecholamine excess, which include paroxysmal episodes of headache, palpitations, sweating, and elevated BP although a significant proportion of patients with PHEOs and non-HN PGL may be diagnosed incidentally. HN PGLs are not classically associated with features related to catecholamine excess and may present as an incidentally detected carotid body/glomus tumor or as a result of pressure effect on surrounding structures.

PHEOs and PGL usually occur sporadically although it is essential to be aware of their association from familial syndromes such as MEN type 2A and 2B (MEN2A and MEN2B) syndrome, von Hippel–Lindau (vHL) syndrome, succinate dehydrogenase (SDH) B, D, and C mutations and neurofibromatosis (NF) type 1 (NF1) [48].

### 13.7.2 Clinical case

A 42-year-old man was referred to an endocrinology clinic for further evaluation of paroxysmal episodes of light headedness, increased sweating, and palpitations. He had a past medical history of seizures and undergone a right-sided nephrectomy, during an age of 26 years, for a renal tumor. He had a family history of hypertension, stroke, and ischemic heart disease. His father had died of an ischemic stroke at an age of 57 despite not having any obvious predisposing factor for stroke apart from hypertension.



On examination, his body mass index was 27.4 kg/m<sup>2</sup>. His pulse rate was 80/minute and he had a BP of 170/86 mm of Hg. His general physical and systemic examinations were unremarkable.

Investigations:

Serum Na	142 mmol/L	(135–145)
Serum K	4.4 mmol/L	(3.5–5.5)
Serum creatinine	94 umol/L	(60–115)
24-hour urinary metadrenaline	Three times as normal	
CT abdomen	A 4.5-cm heterogeneous mass in right adrenal gland with cystic degeneration	

### 13.7.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario

This patient presented with classical clinical features associated with catecholamine excess including paroxysmal episodes of hypertension, increased sweating, and palpitation, although a significant proportion of patients with PHEOs may have sustained hypertension and around 10% of patients may be normotensive. The following clinical characteristics should alert a clinician to screen for PHEOs and/or PGL:

- Hypertension in young patients
- Hypertensive crisis precipitated by general anesthetic agents, nasal decongestants, tricyclic antihypertensive agents
- Malignant or accelerated hypertension
- Unexplained cardiovascular events such as myocardial infarction, pulmonary edema, arrhythmia, shock, and stroke in first-degree relative or index case
- Family history of medullary thyroid cancer (MTC), gastrointestinal stromal tumors, or PHEOs/PGL suggestive of MEN2 syndrome
- Family history of renal carcinoma, hemangioblastoma, pancreatic neuroendocrine tumors (pNET) suggestive of underlying vHL syndrome
- Cyanotic congenital heart disease, erythrocytosis at a young age, or cardiomyopathy
- Incidental adrenal lesion

The past medical history of our patient includes seizures and a nephrectomy for a renal tumor, which suggests a possible underlying diagnosis of vHL syndrome. PHEOs are present in 20%–25% patients with vHL and as a result, these patients need to be on lifelong surveillance even if the initial biochemical screening for metanephries does not yield positive results.

The serum and urinary metanephries results for our patient are consistent with a diagnosis of PHEO, and radiological results (CT adrenal gland) are consistent with this diagnosis. In patients with PGL, most of the tumors are intraabdominal and may occur around periaortic, perinephric, and bladder region although 20% of the patients may have an extraabdominal PGL. The typical site for an extraabdominal PGL includes anterior and posterior mediastinum. Metaiodobenzylguanidine may be an useful investigation in patients with PHEO or PGL if the diagnosis is unclear on initial radiological screening (CT or MRI) or if metastatic disease is suspected.

### 13.7.2.2 Overview of the relevant molecular systems underpinning the clinical scenario

PHEOs and PGLs are usually sporadic although these may be a part of familial syndromes in 15%–20% of the patients. It is important to note that around 30%–40% of patients with PHEOs and PGL have an underlying germline mutation, which may be indicative of variable penetrance of these mutations and/or undiagnosed patients with PHEO and/or PGL in the family in view of variable clinical presentation of this condition, which is considered to be a great mimic. The germline mutations that predispose to PHEO and PGL can be broadly categorized into three groups:

- Cluster 1A, which includes mutations in vHL type 2, Egg-laying gene nine, and endothelial PAS1 gene.
- Cluster 1B, which include mutations involving the tricarboxylic acid cycle (Krebs cycle) and *SDH subunit A*, *subunit B*, *subunit C*, *subunit D*, *malate dehydrogenase*, and *fumarate hydratase*.
- Cluster 3, which includes mutation in RET proto-oncogene, leading to MEN2A and MEN2B syndrome apart from mutations in *NF1* and transmembrane protein 127. Table 13.16 shows the characteristics of various gene mutations associated with PHEOs and PGL [50,51].

**TABLE 13.16** Characteristics of various gene mutations associated with pheochromocytomas (PHEOs) and paraganglioma (PGL).

Gene mutation	Protein encoded	PHEO/PGL	Associated features
<i>vHL2</i>	pVHL proteins involved in degradation of HIF	Early age of onset (mean age <30 years); may be bilateral tumors	Hemangioblastomas Renal cysts RCC Pancreatic cysts and PNET
<i>EGLN</i>	HIF prolyl-hydroxylase	Relatively early age of onset for PHEOs and PGL Recurrent tumor	Congenital erythrocytosis
<i>EPAS1</i>	HIF-2 $\alpha$	Relatively early age of onset for PHEOs Malignant and metastatic	Congenital erythrocytosis Somatostatinoma
<i>SDHA</i>	Subunit A	Sporadic PGL	—
<i>SDHB</i>	Subunit B	PHEOs PGL in sympathetic as well as parasympathetic chain Located from neck to pelvis Malignant and metastatic	Neuroblastoma RCC
<i>SDHC</i>	Subunit C	PGL in sympathetic as well as parasympathetic chain No adrenal PHEOs	
<i>SDHD</i>	Subunit D	Paternaly inherited SDHD-gene patients can develop PGL/PHEOs Parasympathetic HN PGL Malignancy is uncommon	GIST RCC
<i>RET</i>	RET proto-oncogene-activating mutation	PHEOs may be bilateral Mostly benign	MTC Primary HPT in MEN2A Mucosal neuromas and marfanoid features in MEN2B
<i>NF1</i>	NF1 tumor suppressor gene	PHEOs in <5%	Optic glioma, hamartoma, café-au-lait spots

*EGLN*, Egg-laying gene nine; *EPAS1*, endothelial PAS1; *GIST*: gastrointestinal stromal tumor; *HIF*, hypoxia-inducible factors; *HPT*, hyperparathyroidism; *MEN*, multiple endocrine neoplasia; *MTC*, medullary thyroid carcinoma; *NF1*, neurofibromatosis type 1; *PNET*, pancreatic neuroendocrine tumors; *pVHL*, protein encoded by VHL; *RCC*, renal cell carcinoma; *RET*, rearranged during transfection; *SDH*, succinate dehydrogenase; *TMEM*, transmembrane protein 217.

*VHL* gene is a tumor-suppressor gene, which encodes for two different proteins that play a part in degradation of hypoxia-inducible factors (HIF-1 $\alpha$  and HIF-2 $\alpha$ ). HIF is a transcription factor, which induces production of vascular endothelial growth factor, platelet-derived growth factor, glucose transporter, and erythropoietin as an adaptation to hypoxia. Accumulation of HIF promotes the process of tumorigenesis. Germline mutation in *VHL* gene is associated with formation of vascular tumors, such as hemangioblastomas, in the retina and/or cerebellum, renal cell carcinoma, and PHEOs. The PHEOs associated with *VHL* gene mutation may be bilateral and/or malignant and present at a relatively early age. Pancreatic cysts and pNET may be present in 70% and 20% of the patients with *vHL* syndrome, respectively, although majority of these are nonfunctional and detected incidentally.

SDH enzyme is involved in catalyzing the conversion of succinate into fumarate in the mitochondria as part of Krebs cycle. Mutation in subunit A, B, C, D of SDH gene, which encodes for the mitochondrial complex 2, is associated with formation of PGLs. These may be malignant and metastatic although >80% of such tumors are intraabdominal.

**TABLE 13.17** Various syndromes associated with pheochromocytomas (PHEOs) and paragangliomas (PGLs) along with their genetic and clinical characteristics.

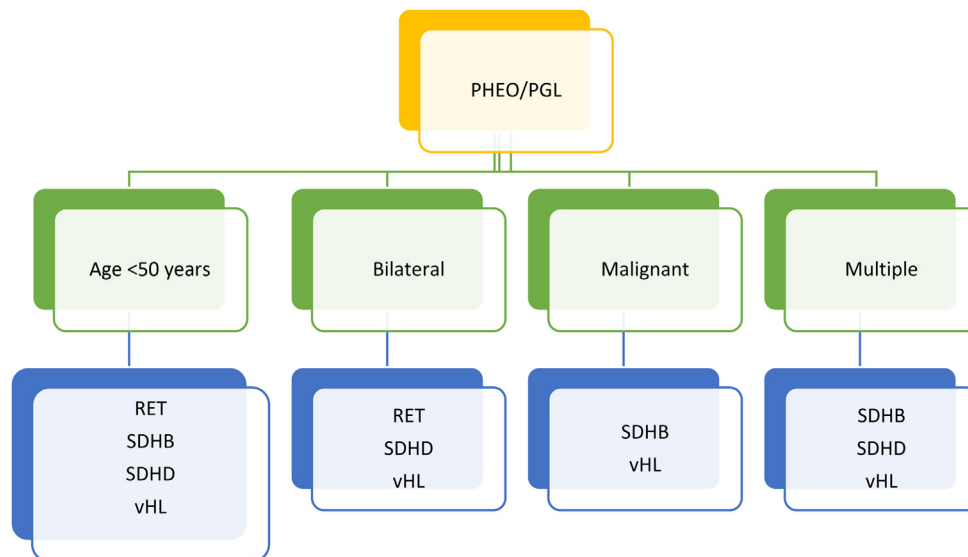
Syndrome	Gene	Biochem	Site	B/L PHEO	Mets	Clinical features
MEN2	<i>RET</i>	Adr	Adrenals	50%–80%	Rare	MTC
vHL	<i>vHL</i>	NA/Adr	Adrenals	50%	<5%	RCC Islet cell tumors Hemangioblastoma
NF1	<i>NF1</i>	Adr	Adrenals	15%	10%	Café-au-lait spots Gliomas
PGL4	<i>SDHB</i>	NA/DA	Symp PGL	Rare	30%–70%	RCC/GIST/Pit adenoma
PGL1	<i>SDHD</i>	NA/DA	HN PGL	Rare	<5%	RCC/GIST/Pit adenoma
PGL3	<i>SDHC</i>	–	HN PGL	–	Rare	RCC/GIST/Pit adenoma
	<i>TMEM 127</i>	Adr/NA	Adr	35%	<5%	Possibly to breast Ca

*Adr*, Adrenaline; *Ca*, carcinoma; *DA*, dopamine; *GIST*, gastrointestinal stromal tumor; *HN PGL*, head and neck paraganglioma; *NA*, noradrenaline; *PGL*, paraganglioma; *Pit*, pituitary; *RCC*, renal cell carcinoma; *Symp*, sympathetic; *vHL*, von Hippel–Lindau.

Germline mutation in *RET* proto-oncogene is associated with the development of MEN2 syndrome. *RET* encodes for a transmembrane receptor tyrosine kinase, which is expressed in the neural crest. MEN2 syndrome is further subdivided into MEN2A, which is characterized by missense mutations in extracellular *RET* domain leading to the constitutive activation of tyrosine kinase. The clinical spectrum in patients with MEN2A syndrome includes development of medullary thyroid carcinoma (MTC) (in 100%), PHEOs (in 40%–50%), and primary HPT (in 20%–30%). The PHEOs tend to occur at a younger age and are more likely to be bilateral. MEN2B syndrome is characterized by a single amino acid substitution (methionine to threonine) in exon 16 of intracellular domain of tyrosine kinase. These patients develop mucosal neuroma (on tongue, lips, eyelids), may have marfanoid features (scoliosis, kyphosis, pectus excavatum), aggressive MTC, and intestinal glioneuromatosis leading to intestinal dysmotility. Primary HPT is not usually associated with MEN2B syndrome.

The genetic mutations, biochemical, clinical, and radiological characteristics of PHEOs and PGL associated with familial germline mutations are shown in Table 13.17 [52,53].

The following algorithm is helpful for genetic testing in patients with suspected PHEO or PGL:



Algorithm for genetic screening in patients with PHEOs and PGL based on clinical and radiological characteristics.

**TABLE 13.18** Suggested protocol for lifetime surveillance for individuals with genetic mutations leading to pheochromocytomas or paraganglioma.

Parameter	<i>vHL2</i>	<i>RET</i>	<i>SDHB, SDHC, and SDHD</i> (paternally inherited)
Clinical assessment	<ul style="list-style-type: none"> <li>Self-monitoring of BP</li> <li>Formal BP measurement every 2 years</li> <li>Yearly retinal examination from age 10–15 years</li> </ul>	<ul style="list-style-type: none"> <li>Self-monitoring of BP</li> <li>Formal BP measurement and neck examination every 2 years</li> </ul>	<ul style="list-style-type: none"> <li>Self-monitoring of BP</li> <li>Formal BP measurement every 2 years</li> </ul>
Biochemical investigations	Yearly plasma/urinary free metanephrines <sup>a</sup> measurement	<ul style="list-style-type: none"> <li>Yearly plasma/urinary free metanephrines</li> <li>Serum calcitonin</li> <li>Serum calcium and PTH (only for MEN2A)</li> </ul>	Yearly plasma/urinary free metanephrines and chromogranin A measurement
Radiological investigations	Every 2 years MRI of brain/spine MRI of abdomen	Yearly thyroid ultrasound if thyroid is intact	Yearly ultrasound of neck, abdomen, and pelvis MRI of head, neck (in SDHC or SDHD) every 3 years
Miscellaneous	In patients with abnormal biochemical screening, MRI or CT of abdomen, or MIBG or FDG-PET for confirmation and evaluation of metastasis	Prophylactic thyroidectomy by the age of 5 years in MEN2A and by 6 months in MEN2B	Consider endoscopy if GIST is suspected

BP, Blood pressure; FDG-PET, 18-Fluoro-2-deoxy-glucose positron emission tomography; GIST, gastrointestinal stromal tumor; MEN2B, multiple endocrine neoplasia type 2B; MIBG, metaiodobenzylguanidine; PTH, parathyroid hormone; RET, rearranged during transfection; *vHL2*, von Hippel–Lindau syndrome 2.

### 13.7.2.3 Management of pheochromocytomas and paraganglioma

Patients with PHEOs are usually treated with elective adrenalectomy. Preoperative management of these patients includes initiation and uptitration of  $\alpha$  blockers such as phenoxybenzamine.  $\beta$ -Blockers should only be introduced once patients are adequately  $\alpha$ -blocked to prevent precipitation of PHEO crisis due to unopposed  $\alpha$  blockade–mediated vasoconstriction. Most adrenal PHEOs <6 cm in size can be removed by laparoscopic surgery although the larger tumors may warrant an open laparotomy.

Patients and relatives harboring gene mutations associated with PGL and PHEOs need lifelong surveillance. Table 13.18 shows the suggested protocol for lifetime surveillance of PHEO and PGL patients based on underlying mutation:

## 13.8 Primary gonadal failure and androgen resistance/insensitivity syndrome

### 13.8.1 Primary gonadal failure

#### 13.8.1.1 Introduction

Primary gonadal failure is a biochemical diagnosis characterized by low testosterone or estradiol levels along with a compensatory increase in gonadotropin levels (raised LH and FSH). It can be congenital or acquired as a result of trauma, irradiation, autoimmunity, or chemotherapy. The commonest etiology leading to primary gonadal failure remains Klinefelter syndrome in men and Turner syndrome in women although infective, infiltrative, and iatrogenic factors can also lead to primary gonadal failure.

#### 13.8.1.2 Clinical case

A 42-year-old man was referred to an endocrinology clinic in view of his low testosterone levels, which were noticed while undergoing investigation in an infertility clinic. He was in a regular relationship for 5 years. He had a past medical history of delayed onset of puberty although there was no history of testicular trauma, mumps orchitis, or systemic illness. He was not on any medication. He denied use of recreational drugs, anabolic steroids, or over-the-counter herbal products.

On examination, his body mass index was 28.5 kg/m<sup>2</sup>. He had bilateral gynecomastia and sparse facial as well as body hair. His testes were small in size and firm in consistency.

## Investigations:

Serum testosterone	4.5 nmol/L	(9–28)
Serum LH	24.0 U/L	(3–8)
FT4	18.5 mU/L	(8–20)
TSH	1.5 mU/L	(0.3–4.5)
Prolactin	115 U/L	(<330)

**13.8.1.2.1 Discussion with reflection on the molecular systems underpinning the clinical scenario**

This young man has biochemical evidence of primary gonadal failure (low testosterone and elevated LH levels). Klinefelter syndrome remains one of the commonest causes of primary gonadal failure with approximate prevalence of 1:660 men. It is an X-linked chromosomal defect with XXY karyotype. Men with Klinefelter syndrome can present with delayed puberty and reduced libido although it is not uncommon for at least 50% of men with Klinefelter syndrome to remain undiagnosed or present at a later stage with infertility or osteoporosis. Klinefelter syndrome is the underlying etiology for 12% of men presenting with azoospermia. There are no phenotypic stigmata of Klinefelter syndrome at birth and signs, such as gynecomastia and reduced facial/body hair, are only noticeable once individuals attain puberty [1]. Low testosterone levels contribute to delayed epiphyseal closure contributing to an increased length of long bones in limbs.

**13.8.1.2.2 Management of Klinefelter syndrome**

The management includes topical or parenteral testosterone therapy. In view of azoospermia, couples with male partners having Klinefelter syndrome usually require in vitro fertilization. These patients have an increased risk of developing breast carcinoma and non-Hodgkin lymphoma. There is an increased predisposition toward developing emphysema, varicose veins, germ-cell tumors (extragonadal), impaired glucose tolerance test, hypothyroidism, early tooth decay, and mitral valve prolapse [54,55].

**13.8.1.2.3 Clinical case**

A 25-year-old man was referred to an endocrine clinic with progressive and gradual decline in libido and erectile dysfunction. He had no past medical history of systemic illness and was not on any regular medication. There was no significant family history of note.

On examination, his body mass index was 30 kg/m<sup>2</sup>. He had a short stature and round facial appearance. He had a deformity at the elbow. On cardiac auscultation, a right-parasternal systolic murmur grade of 3/6 was audible. His testes were small in size and soft in consistency.

Serum testosterone	3.8 nmol/L	(9–28)
Serum LH	31.0 U/L	(3–8)
FT4	14.5 mU/L	(8–20)
TSH	2.3 mU/L	(0.3–4.5)
Prolactin	95 U/L	(<330)

**Discussion with reflection on the molecular systems underpinning the clinical scenario** This young man has features of hypogonadism, which have manifested in peri-pubertal age. Biochemical investigation results are consistent with primary gonadal failure (low testosterone and elevated LH). In addition, he has skeletal features, such as cubitus valgus, short stature, and round face, suggestive of underlying diagnosis of Noonan syndrome, which is either sporadic or familial. It is characterized by the mutation of *PTPN11* gene on chromosome. The karyotype of such men is 46 XY. Such patients may have short stature, webbed neck, bleeding diathesis, cubitus valgus, and right-sided cardiac anomalies [56].

**Androgen resistance/insensitivity syndrome**

**Introduction** Androgen resistance or androgen insensitivity syndrome is a rare disorder of sex development.

**Clinical scenario** A 17-year-old girl was referred to an endocrinology clinic in view of primary amenorrhea. She was a full-term baby born after a normal vaginal delivery. She had no history of any delay in attaining developmental milestones. There was no history of systemic illness although she had an inguinal hernia repair as a child. She was not on any medication. In personal history, she was a nonsmoker, had nil alcohol intake, and was not in any regular relationship.

On examination, her body mass index was 23.5 kg/m<sup>2</sup>. Her BP was 110/76 mm of Hg. She had Tanner stage 4 breasts. Axillary and pubic hair were present. Her general physical and systemic examination were unremarkable.

Investigations:

Serum testosterone	16.9 nmol/L	(<1.5)
Serum LH	14.8 U/L	(3–8)
Serum FSH	4.0 U/L	(1.4–18.1)
Serum prolactin	Normal	
TFT	Normal	
Estradiol	120 pmol/L	(17–260 preludeal)

**Discussion with reflection on the molecular systems underpinning the clinical scenario** Our patient had primary amenorrhea and phenotypic profile of a female although her testosterone levels were significantly elevated for a female patient. Despite having elevated testosterone levels, there were no clinical features, such as excessive hair growth or acne, which are typically associated with hyperandrogenism. Her clinical and biochemical profile was consistent with a possible diagnosis of androgen insensitivity syndrome (AIS), which is an X-linked disorder due to loss of function mutation in the *AR* gene. The karyotype of patients with AIS is XY. AIS can be partial or complete. Individuals with complete AIS have female phenotypic profile although female internal genitalia, such as uterus, fallopian tube, and upper vagina, are absent. Undescended testes may be detected on ultrasound [57].

**Management of patients with androgen insensitivity syndrome** The management of such patients should include psychological support apart from hormone replacement therapy including estrogen and progesterone. As the undescended tests are at a high risk of development of malignancy, a gonadectomy should be considered at the earliest age. The main differential diagnosis of this condition remains 5- $\alpha$  reductase deficiency, which is an autosomal-recessive condition due to mutation of 5- $\alpha$  reductase type 2 (*SRD5A2*). 5- $\alpha$  Reductase mediates conversion of testosterone to its biologically active metabolite dihydrotestosterone (DHT). Male patients with 5- $\alpha$  reductase may present with delayed puberty and low libido. In females, excess testosterone can lead to features of virilization. Patients with a complete absence of this enzyme may have ambiguous genitalia, while partial absence may lead to cryptorchidism and hypospadias. The testosterone levels are high although DHT levels are characteristically low. On radiological assessment, Wolffian duct-derived structures are present although prostate gland remains small. Patients usually require a supraphysiological high dose of testosterone therapy [58].

## References

- [1] Crisafulli G, Aversa T, Zirilli G, De Luca F, Gallizzi R, Wasniewska M. Congenital hypopituitarism: how to select the patients for genetic analyses. *Ital J Pediatr* 2018;44:47.
- [2] Murray PG, Dattani MT, Clayton PE. Controversies in the diagnosis and management of growth hormone deficiency in childhood and adolescence. *Arch Dis Child* 2016;101:96–100.
- [3] Otto AP, França MM, Correa FA, Costalonga EF, Leite CC, Mendonça BB, et al. Frequent development of combined pituitary hormone deficiency in patients initially diagnosed as isolated growth hormone deficiency: a long term follow-up of patients from a single center. *Pituitary* 2015;18(4):561–7.
- [4] Haddad NG, Eugster EA. Hypopituitarism and neurodevelopmental abnormalities in relation to central nervous system structural defects in children with optic nerve hypoplasia. *J Pediatr Endocrinol Metab* 2005;18:853–8.
- [5] De Rienzo F, Mellone S, Bellone S, Babu D, Fusco I, Prodam F, et al. Frequency of genetic defects in combined pituitary hormone deficiency: a systematic review and analysis of a multicenter Italian cohort. *Clin Endocrinol* 2015;83:849–60.
- [6] Giordano M. Genetic causes of isolated and combined pituitary hormone deficiency. *Best Pract Res Clin Endocrinol Metab* 2016;30:679–91.



- [7] Stieg MR, Renner U, Stalla GK, Kopczak A. Advances in understanding hypopituitarism. Version 1. *F1000Res* 2017;6:178.
- [8] Pérez Millán MI, Vishnopolksa SA, Daly AZ, Bustamante JP, Seilicovich A, et al. Next generation sequencing panel based on single molecule molecular inversion probes for detecting genetic variants in children with hypopituitarism. *Mol Genet Genomic Med* 2018;6(4):514–25.
- [9] Bianco SD, Kaiser UB. The genetic and molecular basis of idiopathic hypogonadotrophic hypogonadism. *Nat Rev Endocrinol* 2009;5:569–76.
- [10] Boehm U, Bouloux P, Dattani MT, Roux N, et al. Expert consensus document: European Consensus Statement on congenital hypogonadotropic hypogonadism-pathogenesis, diagnosis and treatment. *Nat Rev Endocrinol* 2015;11:547–64.
- [11] Mitchell AL, Dwyer A, Pitteloud N, Quniton R. Genetic basis and variable phenotypic expression of Kallmann syndrome: towards a unifying theory. *Trends Endocrinol Metab* 2011;22:249–58.
- [12] Herman-Giddens ME, et al. Secondary sexual characteristics in boys: data from the Pediatric Research in Office Settings Network. *Pediatrics* 2012;130:e1058–68.
- [13] Sorensen K, Aksglaede L, Petersen JH, Juul A. Recent changes in pubertal timing in healthy Danish boys: associations with body mass index. *J Clin Endocrinol Metab* 2010;95:263–70.
- [14] Dunkel L, et al. Developmental changes in 24-hour profiles of luteinizing hormone and follicle-stimulating hormone from prepuberty to midstages of puberty in boys. *J Clin Endocrinol Metab* 1992;74:890–7.
- [15] Young J. Approach to the male patient with congenital hypogonadotropic hypogonadism. *J Clin Endocrinol Metab* 2012;97:707–18.
- [16] Dwyer AA, Raivio T, Pitteloud N. Gonadotrophin replacement for induction of fertility in hypogonadal men. *Best Pract Res Clin Endocrinol Metab* 2015;29:910103.
- [17] Raivio T, et al. Reversal of idiopathic hypogonadotrophic hypogonadism. *N Engl J Med* 2007;357:863–73.
- [18] Adami S, Marcocci C, Gatti D. Epidemiology of primary hyperparathyroidism in Europe. *J Bone Miner Res* 2002;17:N18–23.
- [19] Thakker R. Genetics of parathyroid tumours. *J Intern Med* 2016;280:574–83.
- [20] Maruani G, Hertig A, Paillard M, Houillier P. Normocalcaemic primary hyperparathyroidism: evidence for a generalized target-tissue resistance to parathyroid hormone. *J Clin Endocrinol Metab* 2003;88:4641–8.
- [21] Vargas-Poussou R, Mansour-Handili L, Baron S, Bertocchio V, et al. mFamilial hypocalciuric hypercalcaemia types 1 and 3 and primary hyperparathyroidism: similarities and differences. *J Clin Endocrinol Metab* 2016;101:2185–95.
- [22] Marini F, Cianferotti L, Giusti F, Brandi ML. Molecular genetics in primary hyperparathyroidism: the role of genetic tests in differential diagnosis, disease prevention strategy, and therapeutic planning. A 2017 update. *Clin Cases Miner Bone Metab* 2017;14(1):60–70.
- [23] Cardoso L, Stevenson M, Thakker RV. Molecular genetics of syndromic and non-syndromic forms of parathyroid carcinoma. *Hum Mutat* 2017;38(12):1621–48.
- [24] Mizamtsidi M, Nastos C, Mastorakos G, Dina R, Vassiliou I, Gazouli M, et al. Diagnosis, management, histology and genetics of sporadic primary hyperparathyroidism: old knowledge with new tricks. *Endocr Connect* 2018;7(2):R56–68.
- [25] Bradley KJ, Cavaco BM, Bowl MR, Harding B, Cranston T, Fratter C, et al. Parafibromin pathogenic variants in hereditary hyperparathyroidism syndromes and parathyroid tumours. *Clin Endocrinol (Oxf)* 2006;64:299–306.
- [26] Frank-Raue K, Haag C, Schulze E, Keuser R, Raue F, Dralle H, et al. CDC73-related hereditary hyperparathyroidism: five new pathogenic variants and the clinical spectrum. *Eur J Endocrinol* 2011;165:477–83.
- [27] Khan A, Hanley D, Rizzoli R, Bollerslev J, Young J, et al. Primary hyperparathyroidism: review and recommendations on evaluation, diagnosis and management. A Canadian and International Consensus. *Osteoporos Int* 2017;28:1–19.
- [28] Eastell R, Brandi ML, Costa AG, D'Amour P, Shoback DM, Thakker RV. Diagnosis of asymptomatic primary hyperparathyroidism: proceedings of the fourth international workshop. *J Clin Endocrinol Metab* 2014;99(10):3570–9.
- [29] Bilezikian JP, Khan A, Potts JT, Brandi ML, et al. Hypoparathyroidism in the adult: epidemiology, diagnosis, pathophysiology, target-organ involvement, treatment, and challenges for future research. *J Bone Miner Res* 2011;26(10):2317–37.
- [30] Brandi ML, Bilezikian JP, Shoback D, Bouillon R, et al. Management of hypoparathyroidism: summary statement and guidelines. *J Clin Endocrinol Metab* 2016;101:2273.
- [31] Lenos MC, Thakker RV. GNAS mutations in pseudohypoparathyroidism type 1a and related disorders. *Hum Mutat* 2015;36:11.
- [32] Elli FM, deSanctis L, Ceoloni B, Barbieri AM, et al. Pseudohypoparathyroidism Type 1a and pseudo-pseudohypoparathyroidism: the growing spectrum of GNAS inactivating mutations. *Hum Mutat* 2013;34(3):411–16.
- [33] Christov M, Juppner H. Insights from genetic disorders of phosphate homeostasis. *Semin Nephrol* 2013;33:143.
- [34] Chesher D, Oddy M, Darbar U, et al. Outcome of adult patients with X-linked hypophosphatemia caused by PHEX gene mutations. *J Inher Metab Dis* 2018;41:865–76.
- [35] Carpenter TO, Imel EA, Holm IA, Jan de Beur SM, Insogna KL. A clinician's guide to X-linked hypophosphatemia. *J Bone Miner Res* 2011;26(7):1381–8.
- [36] Käyser SC, Dekkers T, Groenewoud HJ, van der Wilt GJ, Carel Bakx J, van der Wel MC, et al. Study heterogeneity and estimation of prevalence of primary aldosteronism: a systematic review and meta-regression analysis. *J Clin Endocrinol Metab* 2016;101:2826–35.
- [37] Monticone S, Burrello J, Tizzani D, Bertello C, Viola A, Buffolo F, et al. Prevalence and clinical manifestations of primary aldosteronism encountered in primary care practice. *J Am Coll Cardiol* 2017;69:1811–20.
- [38] Rossi E, Perazzoli F, Negro A, Magnani A. Diagnostic rate of primary aldosteronism in Emilia-Romagna, Northern Italy, during 16 years (2000-2015). *J Hypertens* 2017;35:1691–7.
- [39] Douma S, Petidis K, Doulas M, Papaefthimiou P, Triantafyllou A, et al. Prevalence of primary hyperaldosteronism in resistant hypertension: a retrospective observational study. *Lancet* 2008;371(9628):1921–6.
- [40] Lenzini L, Rossitto G, Maiolino G, Letizia C, Funder JW, Rossi GP. A meta-analysis of somatic KCNJ5 K<sup>+</sup> channel mutations in 1636 patients with an aldosterone-producing adenoma. *J Clin Endocrinol Metab* 2015;100(8):E1089–95.
- [41] Funder JW. Primary aldosteronism: seismic shifts!. *J Clin Endocrinol Metab* 2015;100:2853–5.
- [42] Auchus R, Arlt W. Approach to the patient: the adult with congenital adrenal hyperplasia. *J Clin Endocrinol Metab* 2013;98(7):2645–55.

- [43] Merke DP, Poppas D. Management of adolescents with congenital adrenal hyperplasia. *Lancet Diabetes Endocrinol* 2013;1(4):341–52.
- [44] Hannah-Shmouni F, Chen W, Merke DP. Genetics of congenital adrenal hyperplasia. *Best Pract Res Clin Endocrinol Metab* 2009;23(2):181–92.
- [45] Auchus RJ. Steroid 17-hydroxylase and 17,20-lyase deficiencies, genetic and pharmacologic. *J Steroid Biochem Mol Biol* 2017;165(Pt A):71–8.
- [46] Yau M, Haider S, Khattab A, Ling C, et al. Clinical, genetic, and structural basis of apparent mineralocorticoid excess due to 11 $\beta$ -hydroxysteroid dehydrogenase type 2 deficiency. *Proc Natl Acad Sci USA* 2017;114(52):E11248–56.
- [47] Lin S, Yang S, Chau T. A practical approach to genetic hypokalemia. *Electrolyte Blood Press* 2010;8(1):38–50.
- [48] Lenders JWM, Duh QY, Eisonhofer G, Gimenez-Roqueplo A, et al. Pheochromocytoma and paraganglioma: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab* 2014;99(6):1915–42.
- [49] McNeil AR, Blok BH, Burke MP, Hilton JM. Pheochromocytomas discovered during coronal autopsies in Sydney, Melbourne and Auckland. *Aust N Z J Med* 2000;30:648–52.
- [50] Neumann HP, Bausch B, McWhinney SR, Bender BU. Germ-line mutations in nonsyndromic pheochromocytoma. *N Engl J Med* 2002;346(19):1459–66.
- [51] Neumann HP, Young Jr WF, Krauss T, Bayley JP, et al. 65 Years of the double helix: genetics informs precision practice in the diagnosis and management of pheochromocytoma. *Endocr Relat Cancer* 2018;25(8):T201–19.
- [52] Toledo RA, Burnichon N, Cascon A, Benn DE, et al. Consensus statement on next-generation-sequencing-based diagnostic testing of hereditary pheochromocytomas and paragangliomas. *Nat Rev Endocrinol* 2017;13:233–47.
- [53] Fishbein L, Leshchiner I, Walter V, Danilova L, et al. Comprehensive molecular characterization of pheochromocytoma and paraganglioma. *Cancer Cell* 2017;31:181–93.
- [54] Groth KA, Skakkebaek A, Høst C, Gravholt CH, Bojesen A. Clinical review: Klinefelter syndrome—a clinical update. *J Clin Endocrinol Metab* 2013;98(1):20–30. Available from: <https://doi.org/10.1210/jc.2012-2382> Epub 2012 Nov 1.
- [55] Chang S, Skakkebaek A, Gravholt CH. Klinefelter syndrome and medical treatment: hypogonadism and beyond. *Hormones* 2015;14(4):531–48.
- [56] Allanson JE, Roberts AE. Noonan syndrome. *Gene Reviews*. Seattle, WA: University of Washington; 2001 [Updated 25.02.16].
- [57] Batista RL, Costa EMF, Rodrigues AS, Gomes NL, et al. Androgen insensitivity syndrome: a review. *Arch Endocrinol Metab* 2018;62(2):227–35.
- [58] Viswanathan V, Eugster EA. Etiology and treatment of hypogonadism in adolescents. *Pediatr Clin North Am* 2011;58(5):1181–200 x.

# Genetic disorders of lipoprotein metabolism

*Deepak Bhatnagar<sup>1</sup>, Jonathan Schofield<sup>1,2</sup> and Handrean Soran<sup>1</sup>*

<sup>1</sup>Cardiovascular Research Group, Division of Diabetes, Endocrinology & Gastroenterology, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom <sup>2</sup>Manchester Diabetes Centre, Manchester Royal Infirmary, Manchester University Hospitals NHS Foundation Trust, Manchester, United Kingdom

## 14.1 Introduction

Lipids, along with proteins and carbohydrates, play a major role in metabolism. They are defined as oils, waxes, fatty acids, and cholesterol. Lipids have a particular role in the physiology of energy metabolism and in the formation of steroid hormones. They are also involved in intracellular signaling, in modulating nerve function, and are responsible for the structural integrity of the cell membrane. In addition, they regulate processes such as protein modification and inflammation. To enable such a diverse range of functions, the pathways of lipid metabolism are closely regulated and also integrated with nonlipid pathways [1]. In order to transport lipids through a predominantly aqueous environment in the body, they are complexed with proteins called apolipoproteins to form macromolecular structures called lipoproteins (Table 14.1) [2]. The apolipoproteins also function as ligands to enable the interaction between lipoproteins and their receptors (Table 14.1) [2] and play a role as inhibitors or cofactors for enzymes involved in lipid metabolism (Table 14.1) [2].

Lipoproteins can be classified according to their molecular mass or molecular diameter or hydrated density or composition. However, by convention, the terminology of lipoproteins is based on their hydrated density, for example, the term low-density lipoprotein (LDL) (Fig. 14.1) [3]. This chapter will deal mainly with lipoprotein metabolism. The terminology of lipoproteins and their function is described in Table 14.1.

## 14.2 An outline of lipoprotein metabolism

In brief, cholesterol, fatty acids, and fat soluble vitamins from food are absorbed in the intestine. In the enterocyte the cholesterol gets esterified to form cholesteryl esters (Fig. 14.2) [4]. The longer chain fatty acids, esterified cholesterol, fat soluble vitamins, and other lipids such as phospholipids are packaged along with apolipoprotein B<sub>48</sub> (apoB<sub>48</sub>), apolipoprotein C, and apolipoprotein E to form a particle called a chylomicron. This lipoprotein is then transported via the lymphatic channels to the circulation where it comes in contact with lipoprotein lipase. This enzyme hydrolyzes the triglycerides to form fatty acids that are used for energy, any excess being stored as triglycerides predominantly in the adipocytes, but to a certain extent also in the liver. Thus the chylomicron particle becomes smaller in hydrated density, but the cholesterol, phospholipids, and apolipoproteins remain in the particle. This chylomicron remnant is taken up by the liver through a putative remnant receptor, delivering the cholesterol, and the fat soluble vitamins to the liver.

Just as the gut produces triglyceride-rich lipoproteins, so does the liver. This time the lipoprotein is called very low-density lipoprotein (VLDL), and it contains cholesterol, triglycerides, phospholipids, and vitamin derived from the liver. These lipids are combined with apolipoprotein B<sub>100</sub> (apoB<sub>100</sub>). On entering the plasma, VLDL acquires apolipoprotein E and C through a series of complex exchanges and transfers. As with the

**TABLE 14.1** The terminology of lipoproteins, apolipoproteins, and associated enzymes.

Lipoproteins	Origin	Function in lipoprotein metabolism
Chylomicrons	Intestine	Transport of dietary triglycerides and cholesterol
Chylomicron remnants	Product of chylomicrons	Delivery of dietary triglycerides and cholesterol to liver
VLDLs	Liver	Transport of endogenous triglycerides and cholesterol
IDL	Product of VLDL	Part of delivery pathway for cholesterol transport to tissues
LDLs	Product of IDL	Delivery of cholesterol to tissues
HDLs	Intestine and liver	Transport of cholesterol from the tissues back to the liver
Lp(a)	Liver	LDL-like particle with apolipoprotein(a), which has homology with plasminogen. Function unclear
Apolipoproteins	Associated with which lipoprotein	Function in lipoprotein metabolism
AI	HDL, chylomicrons	Activation of LCAT
AII	HDL, chylomicrons	? role
AIV	Chylomicrons	? role
B48	Chylomicrons	Assembly of chylomicrons
B100	VLDL, IDL, LDL	Assembly of VLDL. Ligand for LDL receptor
CI	HDL, chylomicrons, VLDL	? role
CII	HDL, chylomicrons, VLDL	Activation of lipoprotein lipase
CIII	HDL, chylomicrons, VLDL, IDL	? role
E	HDL, chylomicrons, VLDL, IDL	Ligand for chylomicron remnant receptor
Enzymes involved in lipoprotein metabolism	Location	Function in lipoprotein metabolism
LCAT <sup>a</sup>	Plasma	Esterification of free cholesterol to cholesteryl esters
CETP <sup>b</sup>	Plasma	Transfer of cholesteryl esters from HDL to LDL and VLDL
Lipoprotein lipase	Capillary endothelium in adipose tissue and muscle	Hydrolysis of chylomicron and VLDL triglycerides
Hepatic lipase	Liver	Hydrolysis of triglycerides in HDL and? IDL

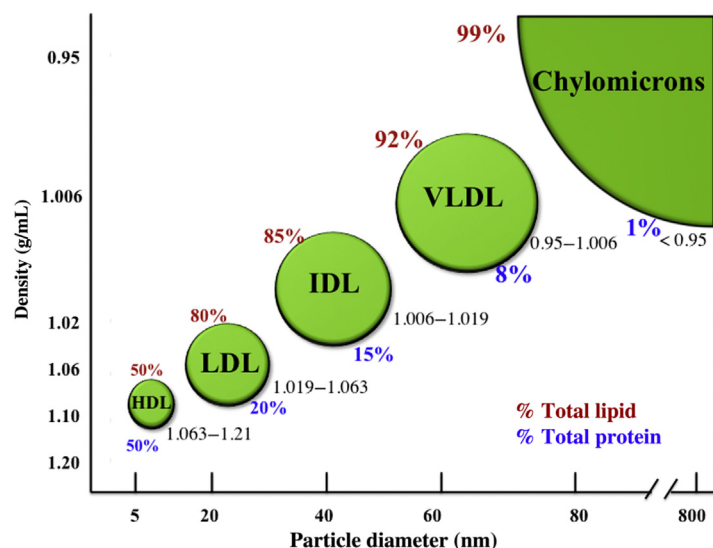
<sup>a</sup>Lecithin—cholesterol acyltransferase.<sup>b</sup>Cholesteryl ester transfer protein.

CETP, Cholesteryl ester transfer protein; HDL, high-density lipoprotein; IDL, intermediate-density lipoprotein; LCAT, lecithin—cholesterol acyltransferase; LDL, low-density lipoprotein; VLDL, very low-density lipoprotein.

With permission from Bhatnagar D. Lipid-lowering drugs in the management of hyperlipidemia. *Pharmacol Ther* 1998;79:205–30.

chylomicron particle, the triglyceride in VLDL too is hydrolyzed by capillary-associated lipoprotein lipase, resulting in a particle of higher density that is termed intermediate-density lipoprotein (IDL). As the triglyceride from the IDL is further hydrolyzed, it results in the formation of LDL. The apolipoprotein B<sub>100</sub> in the LDL particle acts as a ligand for the nearly ubiquitous and very well characterized LDL receptor (LDLR) on the cell surface. This enables the endogenous cholesterol to be transported to the cell.

The body also has mechanisms to deal with excess cholesterol [5]. The main lipoprotein involved in this process is high-density lipoprotein (HDL), which appears to arise from various sources such as the liver, the



**FIGURE 14.1** Lipoprotein classes. The classification of the major types of lipoproteins is based on their densities obtained by flotation ultracentrifugation analysis. The density range for each class is shown, in addition to the lipid (red) and protein (blue) content. The diagram is not to scale. Source: From Jairam V, Uchida K, Narayanaswami V. In: Frank S, Kostner G, editors. *Pathophysiology of lipoprotein oxidation, lipoproteins – role in health and diseases*. IntechOpen; 2012. doi:10.5772/50622. Available from: <https://www.intechopen.com/books/lipoproteins-role-in-health-and-diseases/pathophysiology-of-lipoprotein-oxidation>.

intestine, and exchange of lipids and apolipoproteins in the plasma and interstitial space. The complexity makes it difficult to define the precise metabolism of HDL which is found as several particles that can broadly termed as nascent or as mature after the acquisition and exchange of various apolipoproteins and lipids. HDL is triglyceride poor and rich in cholesterol and cholesteryl esters. HDL helps transport cholesterol from the peripheral tissues to the liver. The protein cholesteryl ester transfer protein (CETP) and the enzyme lecithin–cholesterol acyltransferase (LCAT) play a key role in modeling HDL particles. The other ways the body is able to remove excess cholesterol is by the liver secreting it as bile, the enterocytes secreting cholesterol in the lumen, or the shedding of cells from the skin.

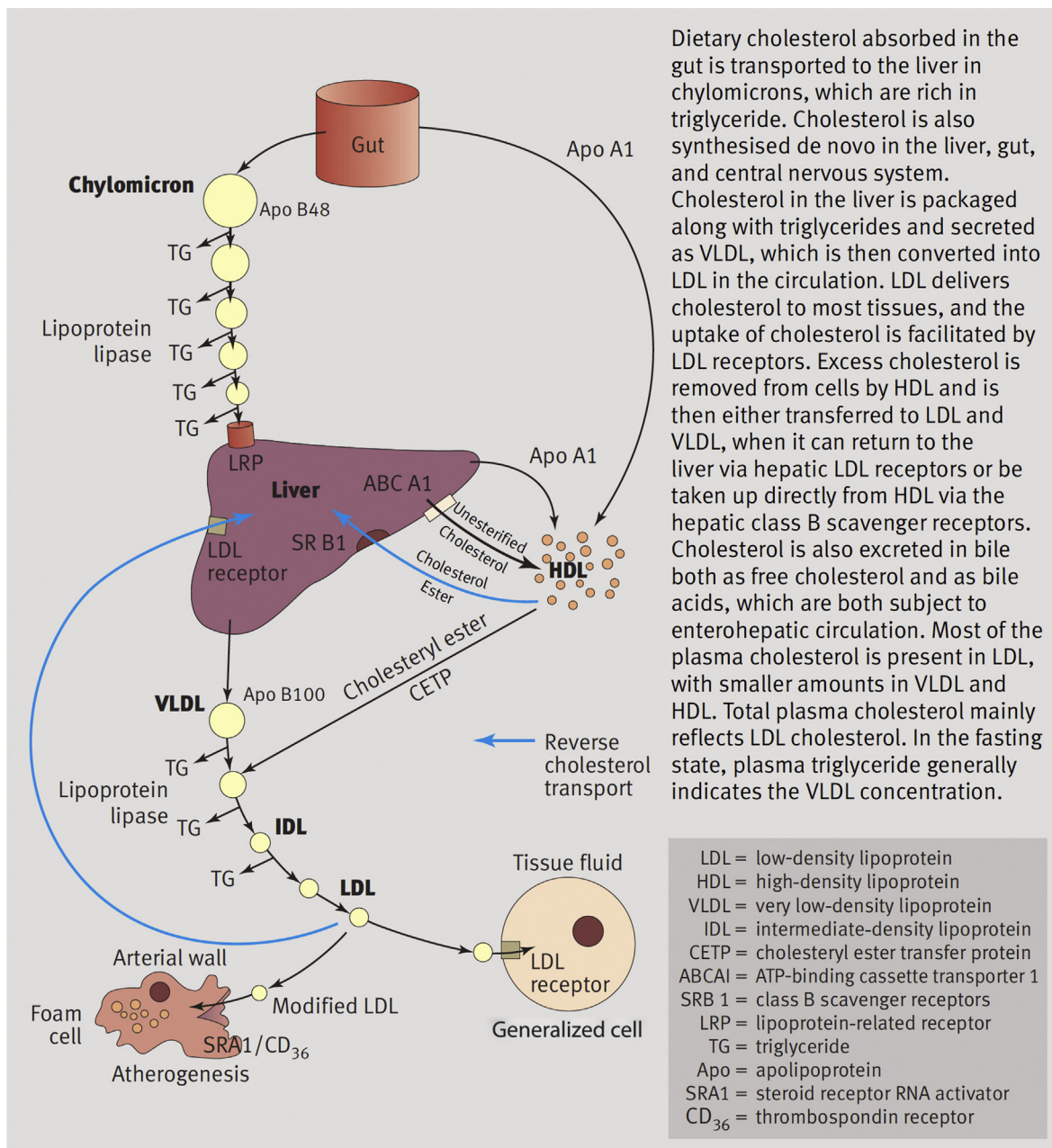
### 14.3 The environmental and genetic factors affecting lipid metabolism

Serum lipid concentrations are predominantly affected by nutritional factors and vary among populations around the world. For example, serum cholesterol concentrations in rural farming communities, especially in the Far East, are particularly low [6–8], but in urban dwellers and in affluent societies, there is a tendency for higher serum cholesterol concentrations [9]. Increasing globalization and improved food security in many parts of have led to higher serum lipid concentrations, particularly in developing countries [10–12] and is largely a result of increase in energy intake and over nutrition. Migration also has an influence of serum lipid levels [13–15]. This association of increasing serum cholesterol concentrations with globalization often generates debate around diet, affluence, and increasing urbanization or “Westernization” and poses a challenge in defining reference ranges. For most analytes the upper limit of “normal” is based on the 95th or 90th centile for a healthy population [16]. Several observational studies and treatment trials indicate that higher serum or LDL cholesterol (LDL-C) concentrations are associated with a higher incidence and prevalence of coronary heart disease (CHD) [17–21]. Given the epidemiological relationship between plasma cholesterol and risk of CHD, this does not apply to plasma cholesterol. There also is no obvious threshold below which this relationship ceases to exist [22]. For this reason, it is more appropriate to base a desirable or healthy concentration of plasma cholesterol on the value at which coronary risk is considered unacceptably high. Most patients developing CHD have plasma cholesterol concentrations that are likely to be between the 30th and 90th centile for their population of origin [23].

Serum triglycerides, the other serum lipid seen on laboratory reports, are also influenced by lifestyle. In many developed societies and in emerging economies, mild hypertriglyceridemia (up to 3–4 mmol/L) is seen frequently and is associated with obesity and, where culturally permitted, increased alcohol intake [15,24].

Apart from lifestyle, inherent biological variation also affects serum lipid concentrations. It is normal for serum triglycerides to increase after a meal. However, serum cholesterol and HDL cholesterol (HDL-C) are unaffected by fasting. Other examples of biological variation include the change in serum lipids with the menstrual cycle. In most societies, serum lipids increase with age, and there appears to be a seasonal variation in temperate climates. From twin and other genetic studies, it has been estimated that around 40% of LDL-C is inherited [25].





**FIGURE 14.2** The role of cholesterol in the metabolism of lipoprotein. Source: Adapted from Charlton-Menys V, Durrington PN. Human cholesterol metabolism and therapeutic molecules. *Exp Physiol* 2008;93(1):27–42. doi:10.1113/expphysiol.2007.035147. With permission from Bhatnagar D, Soran H, Durrington PN. Hypercholesterolaemia and its management. *BMJ* 2008;337:a993.

The study of genetic lipoprotein disorders uncovering monogenic mutations has partially helped to understand how genes influence lipoprotein metabolism [26]. The fall in the cost of genotyping and development of high throughput techniques has led to the investigation of the association between genotypes and lipid phenotypes. These genome-wide association studies (GWAS) have mainly been done in people with normal serum lipids linking single-nucleotide polymorphisms to lipid traits and/or cardiovascular risk [27]. Combining this method with meta-analysis in studying large-scale populations such as participants in bio banks and other groups has identified a large number of both common and rare genetic variants [28]. However, the size of the effect of these genetic variants only explains less than 20% of the phenotypic variation in certain lipids [29,30]. To overcome this low size effect of the genetic variants and use these data clinically, many groups have tried to



develop genetic risk scores based on the GWAS loci [31]. These scores are not yet ready for clinical use but may improve with the development of new microarrays.

### 14.4 Screening for lipoprotein disorders

Given the general global rise in serum lipid levels, it is advisable to screen the whole population for dyslipidemia as part of cardiovascular disease (CVD) risk assessment from the age of 40 years [32–34]. This is commonly done using risk scores, but they should not be used where cardiovascular risk is deemed already high enough to warrant lifestyle advice and other cardiovascular treatment. Examples include patients with atherosclerotic CVD, hypertension with target organ damage, patients with type 1 or type 2 diabetes mellitus, and renal dysfunction (including diabetic nephropathy). In addition, risk scores should not be used in patients with primary hyperlipidemia, such as familial hypercholesterolemia (FH), familial combined hyperlipidemia (FCH), or other inherited dyslipidemia. As in all clinical matters, assessment and treatment should be guided by the benefits and risks of treatment, informed preference, and comorbidities that may make treatment inappropriate. There are several risk assessment tools used around the world. These are based on methods using multivariate mathematical models, which differ according to the risk factors they incorporate. In the United Kingdom, National Institute for Health and Clinical Excellence (NICE) [32] has left practitioners to choose from Joint British Societies' consensus recommendations [JBS3 (<http://www.jbs3risk.com>)] [34] or QRISK Lifetime (<http://www.qrisk.org/lifetime/>) [QRISK2 (an algorithm based on an amalgamated UK general practice database) (<http://www.qrisk.org/>)] [35], or QRISK3 (<https://www.qrisk.org/three>) or ASSIGN (an algorithm based on an amalgam of Scottish epidemiological studies, the Scottish Heart Extended Cohort) (<http://assign-score.com/>).

In Europe the systematic coronary risk estimation (SCORE) [36] charts are recommended, which estimate the 10-year risk of fatal CVD (CHD, stroke, or other occlusive arterial disease) as rate per 100% over 10 years. In the United States a points system based on the Framingham Heart Study is employed to give the 10-year likelihood of CHD (fatal and nonfatal myocardial infarction) as rate per 100% over 10 years [37]. The 2013 American College of Cardiology/American Heart Association (ACC/AHA) Task Force developed risk equations for non-Hispanic African-American and White men and women aged 40–79 years. The 10-year risk estimation was based on nonfatal myocardial infarction, CHD death, fatal, and nonfatal stroke [38]. It is important to note that the risk calculation methods can only be used with the guideline for which they were designed [39]. An alternative more practical method is to base the decision of prescribing lipid-lowering therapy for primary prevention on beneficiaries achieved, derived from number needed to treat or cardiovascular events prevented per 100 patients treated over a period of time, for example, 10 years [40–42].

### 14.5 Diagnosing genetic disorders of lipoprotein metabolism

Most patients with primary disorders of lipoprotein metabolism are diagnosed incidentally when clinicians come across very high serum lipid concentrations or those with premature CHD. When challenged by a persistently poor lifestyle and diet, the body is generally able to limit the rise in serum lipids. However, depending on background population serum lipid levels, a serum cholesterol levels higher than 7.0 or 7.5 mmol/L or moderate-to-severe hypertriglyceridemia should raise the suspicion of a genetic disorder of lipoprotein metabolism. There are also certain signs (Table 14.2) [2] that point to the presence of a genetic disorder. However, it is important to note that in some secondary hyperlipidemias, serum lipid concentrations may be as high as those seen in primary hyperlipidemia. The coexisting secondary disorder is generally clinically apparent in such cases (Table 14.3) but may require other tests to establish the secondary cause.

An etiological classification of primary lipoprotein disorders is presented in Table 14.4 [2]. A more commonly used and clinical classification of frequently encountered primary hyperlipidemias is shown in Table 14.5 [16].

### 14.6 Common (polygenic) hypercholesterolemia

The commonest cause of hypercholesterolemia is overproduction of VLDL leading to increased LDL [1]. Polymorphic variants within certain genes may influence lipoprotein production or clearance in the presence of nutritional excess relative to energy expenditure [43,44]. If the gene variant does not affect the conversion of

**TABLE 14.2** Some signs associated with disorders of lipoprotein metabolism.**Eyes**

Corneal arcus

Corneal opacities

**Skin**

Xanthelasmata

Tendon xanthomata

Eruptive xanthomata

Tuberoeruptive xanthomata

Palmar xanthomata

Planar xanthomata

**Systemic**

Aortic stenosis

Hepatosplenomegaly

Lipemia retinalis

Peripheral neuropathy

Ataxia

*With permission from Bhatnagar D. Lipid-lowering drugs in the management of hyperlipidaemia. Pharmacol Ther 1998;79:205–30.*

**TABLE 14.3** Exclude underlying causes of hyperlipidemia before diagnosing primary hypercholesterolemia.**Some causes of secondary hyperlipidemia**

Nephrotic syndrome

Obstructive jaundice

Hypothyroidism

Cushing's syndrome

Anorexia nervosa

Thiazide diuretics

**CiclosporinSome causes of mixed hyperlipidemis<sup>a</sup>**

Type 2 diabetes mellitus

Obesity, particularly if accompanied by features of the metabolic syndrome

Alcohol excess

Monoclonal gammopathy

Renal dialysis

Glucocorticoids

β-Blockers

Retinoic acid derivatives

<sup>a</sup>Increase in plasma cholesterol and triglycerides.

*With permission from Bhatnagar D, Soran H, Durrington PN. Hypercholesterolaemia and its management. BMJ 2008;337:a993.*

**TABLE 14.4** An etiological classification of primary lipoprotein disorders.

Disorder	Remarks
<b>Disorders due to defects in lipoprotein receptors</b>	
Familial hypercholesterolemia	Serum cholesterol elevated since birth premature CAD defect in LDL receptor
<b>Disorders due to overproduction of lipoproteins</b>	
Familial combined hyperlipidemia	Mixed hyperlipidemia, usually apparent around 30–40 years of age, no characteristic clinical signs, premature CAD, overproduction of VLDL
Familial hypertriglyceridemia	Moderate-to-severe hypertriglyceridemia, risk of pancreatitis, overproduction of VLDL
<b>Disorders due to defects in apolipoproteins</b>	
Remnant hyperlipidemia	Mixed hyperlipidemia, characteristic clinical signs, premature CHD, and PVD apolipoprotein E2 homozygosity
Familial defective apolipoprotein B	Biochemical and clinical features similar to heterozygous FH, defective ligand for LDL receptor
Apolipoprotein CII deficiency	Severe hypertriglyceridemia due to lack of lipoprotein lipase activity
Familial hypoalphalipoproteinemia	Low HDL cholesterol values
Abetalipoproteinemia	Very low serum triglycerides and cholesterol, progressive neurological disorder, steatorrhea
<b>Disorders due to defects of enzymes and proteins involved in lipoprotein metabolism</b>	
Familial hyperchylomicronemia	Severe hypertriglyceridemia due to lack of lipoprotein lipase
Hepatic lipase deficiency	Predominant hypertriglyceridemia, clinical signs similar to remnant hyperlipidemia
$\alpha$ -Lecithin–cholesterol acyltransferase deficiency	Hypertriglyceridemia, proteinuria, corneal opacities
Cholesteryl ester hydrolase deficiency	Hypercholesterolemia, hepatosplenomegaly
Cholesteryl ester transfer protein deficiency	High HDL cholesterol
<b>Miscellaneous disorders</b>	
Polygenic hypercholesterolemia	Hypercholesterolemia associated with a family history
Hyperalphalipoproteinemia	High HDL-cholesterol levels; may be related to CETP deficiency
Hyperapobetalipoproteinemia	Elevated apolipoprotein B in the presence of normal LDL cholesterol values; associated with increased CHD risk
Elevated Lp(a)	Partial homology with plasminogen; associated with increased CHD risk

CAD, Coronary artery disease; CETP, cholesteryl ester transfer protein; FH, familial hypercholesterolemia; HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very low-density lipoprotein.

With permission from Bhatnagar D. Lipid-lowering drugs in the management of hyperlipidaemia. *Pharmacol Ther* 1998;79:205–30.

VLDL to LDL, LDL-C alone will be raised, especially if the variant influences LDL catabolism also. If the conversion of VLDL to LDL is affected by the variant then a rise in VLDL is manifested as a rise in serum triglycerides [1]. Hypertriglyceridemia in the presence of increased LDL-C further increases the risk of CVD [45–48]. If relatives of a patient with a combined increase in LDL-C and triglycerides are screened, some will have principally hypertriglyceridemia, some raised LDL-C alone, others a combined increase whilst some will have relatively normal lipids, depending on the particular combination of polygenic characteristics each has inherited. This is termed FCH [1,33]. Often hypertriglyceridemia is associated with central obesity and insulin resistance and clusters with other risk factors such as low levels of HDL-C, hypertension, and impaired glucose tolerance or type 2 diabetes mellitus [49]. The concept of *metabolic syndrome* recognizes the clustering of CVD risk factors and is defined by the presence of three or more factors such as increased waist circumference, fasting hypertriglyceridemia, a low HDL-C, hypertension or impaired fasting glucose, or use of medication for type 2 diabetes mellitus [50]. CVD risk is increased in this prodromal period, often years before diabetes develops [51]. There is clearly overlap between FCH, metabolic syndrome, and type 2 diabetes mellitus [52–54].

**TABLE 14.5** Most commonly encountered causes of primary hypercholesterolemia.

Diagnosis	Prevalence (%)	Inheritance	Clinical features	Biochemistry
Common hypercholesterolemia <sup>a</sup>	70	Polygenic	Usually none (sometimes corneal arcus, xanthelasmata)	Raised cholesterol owing to low-density lipoprotein
Heterozygous familial hypercholesterolemia	0.2	Monogenic	Cholesterol raised from childhood. In adulthood often tendon xanthomata or Achilles tenosynovitis; sometimes corneal arcus and xanthelasmata	Raised cholesterol owing to low-density lipoprotein
Familial defective apolipoprotein B	0.2	Monogenic	Usually none. Occasionally familial hypercholesterolemia phenotype	Raised cholesterol owing to low-density lipoprotein
Combined hyperlipidemia <sup>b</sup>	10	Polygenic	Usually none (sometimes corneal arcus, xanthelasmata); overlap with dyslipidemia of type 2 diabetes and metabolic syndrome	Raised triglycerides and cholesterol owing increased very low-density lipoprotein
Type 3 hyperlipoproteinemia (dysbetalipoproteinemia; remnant removal disease)	0.02	Monogenic	Striate palmar xanthomata; tuberous xanthomata	Raised triglycerides and cholesterol owing to intermediate-density lipoprotein and chylomicron remnants. Usually apo E2 homozygosity
Severe hypertriglyceridemia (> 10 mmol/L)	0.1	Polygenic or monogenic	Eruptive xanthomata; acute pancreatitis; milky plasma; lipemia retinalis	Raised triglycerides owing to fasting chylomicronemia and increased very low-density lipoprotein

<sup>a</sup>Defined as plasma cholesterol  $\geq 5$  mmol/L in middle age.

<sup>b</sup>Defined as plasma cholesterol  $\geq 5$  mmol/L and fasting plasma triglycerides  $\geq 1.7$  mmol/L in the absence of diagnostic features of heterozygous familial hypercholesterolemia (Table 14.1).

Prevalence is approximate and refers to the adult population in the United Kingdom (Durrington, see "Additional education resources" box).

With permission from Bhatnagar D, Soran H, Durrington PN. Hypercholesterolaemia and its management. *BMJ* 2008;337:a993.

## 14.7 Familial hypercholesterolemia

FH is an autosomal dominant inherited disorder of lipoprotein metabolism, characterized by increased LDL-C that is present since birth [55]. If untreated, it leads to premature CHD [56]. It is distinct from simple heritable conditions that lead to increased serum cholesterol. It results from a defect in the LDLR gene [57], which is situated on the short arm of chromosome 19 [58,59]. Its product, the LDLR, is present on most tissues and is responsible for cellular uptake of apoB and apoE-containing lipoproteins. LDLRs on the liver account for clearing the vast majority of LDL particles from the plasma. The recycling and degradation of LDLRs is regulated by the peptide PCSK9. Several hundred mutations in the LDLR gene have been described worldwide and are recorded on an online database <http://www.lovnd.nl/LDLR> (accessed June 3, 2019). There is geographical variation of the mutations, except in populations that have been genetically isolated [60]. Other single-gene mutations causing hypercholesterolemia can also produce a clinical phenotype similar to FH (Table 14.6) [61].

## 14.8 Characteristic clinical features of familial hypercholesterolemia

The hypercholesterolemia of FH results from an increase in LDL particle numbers, resulting from decreased catabolism due to an LDLR defect [62], but there may also be some overproduction of LDL [63]. HDL-C levels are generally low in FH patients, and lipoprotein(a) levels are increased [64,65]. The FH clinical phenotype is variable and is affected by environmental, metabolic, and other genetic factors [66,67].

Patients with heterozygous FH often have tendon xanthomata present in the extensor tendons, typically found in Achilles tendons or the dorsum of the hands [1]. They can be unilateral, and generally appear after the age of 30, although there is some evidence that they appear at an earlier age in FH due to certain mutations. In some patients, Achilles tendinitis can be a presenting symptom, especially so in patients with ill-fitting shoes or those who are physically active. Xanthomata in the Achilles tendon can weaken the tendon sometimes leading to

**TABLE 14.6** Monogenic disorders that can present with a lipid phenotype similar to familial hypercholesterolemia.

Disorder	Gene involved	Inheritance
Classical familial hypercholesterolemia	LDL receptor	Codominant
Familial defective apolipoprotein B	Apolipoprotein B	Codominant
Autosomal dominant familial hypercholesterolemia or FH3 (NARC 1)	PCSK9	Dominant
ARH	ARH (involved in internalization of LDL receptor)	Recessive
Cholesterol 7 alpha hydroxylase deficiency	CYP7A1 (involved in bile acid synthesis)	Recessive
Sitosterolemia	ABC G5 and ABC G8	Recessive

ARH, Autosomal recessive hypercholesterolemia; LDL, low-density lipoprotein; PCSK9, protein convertase subtilisin/kexin 9.

With permission from Bhatnagar D. Health implications of familial hyperlipidemia Chapter 24. In: Kumar D, editor. Genomics and health in the developing world. Oxford University Press; 2012. Available from: <https://doi.org/10.1093/med/9780195374759.003.0024>.

rupture. Most patients will also have corneal arcus, but this can be seen in other genetic lipid disorders, too. Cutaneous xanthelasmata, especially around the eyes, are also nonspecific but are frequently encountered in FH. Aortic stenosis is also seen in patients with heterozygous FH and is nearly always present in patients with homozygous FH [68].

If untreated, patients with heterozygous FH have premature CHD and peripheral vascular disease [69–73], but the relationship between ischemic stroke and FH is unclear [74]. In patients with homozygous FH, sudden death from CHD is not uncommon and often, despite aggressive management with LDL apheresis (lipoprotein apheresis) and prophylactic coronary artery bypass grafting, the prognosis remains poor [69].

The natural history of CHD in FH is drawn from previous studies carried out in the United States, United Kingdom, Scandinavia, Canada, France, Holland, and Japan [75–78]. Not all these studies were well designed, and in some cases, patients were on nonstatin lipid-lowering drugs. The association between the severity of CHD and specific FH mutations has been sought [79–81] but is difficult to establish, as the series of patients examined have been small, with large allelic variation and a variation in the definition of CHD.

More severe form of homozygous FH results when a double dose of the defective gene is inherited, or the subject is a compound heterozygote. These patients invariably develop CHD before the age of 25 years with homozygous FH, and in addition to tendon xanthomata, they can develop cutaneous and tuberous xanthomata on the knees, elbows, buttocks, and the dorsum of the hands. In many cases, CHD appears in childhood but is difficult to diagnose unless there is a high degree of suspicion [69,82].

Hypertension, hypertriglyceridemia, or hyperuricemia, which are features of FCH, are uncommon in patients with FH [1]. The reason for this is not clear, but given the rise in obesity, they can coexist in patients FH and cause diagnostic confusion.

The FH clinical phenotype is variable. People with the same LDLR mutation can have varying degrees of hypercholesterolemia, although generally patients from the same family have similar serum cholesterol levels [83]. Not all the classical features of FH originally described are present in all patients. Genetic, metabolic, and environmental factors can all influence the phenotypic variability, but the studies to establish these data have small numbers and are derived mainly from case–control studies. Attempts have been made to determine the effects of mutations on phenotypic variability by pooling details from different, mainly European, populations. The results are variable but appear to suggest that receptor-negative mutations are associated with a more severe phenotype, compared with receptor-defective mutations [84–86].

The FH phenotype can be affected by genes that influence the assembly or secretion or remodeling of lipoproteins, but it is unusual for them to exert a major influence on the FH phenotype [61]. Gender does seem to have a major influence on the FH phenotype, with women demonstrating clinical features of CHD later than men, but earlier than women without FH. The expression of LDLRs can be influenced by variation in physiological or pathological factors in thyroid hormone or estrogen levels.

Environmental factors such as diet, behavior, and cultural factors also affect the FH phenotype. For example, LDL-C levels and CHD are reported to be lower in FH heterozygotes in China [87,88]. With increasing recognition of elevated serum cholesterol as a risk factor, physicians are increasingly treating hypercholesterolemia at an earlier stage, often without establishing its origin or nature. Therefore statin treatment often modifies the FH

phenotype, delaying the appearance of signs such as tendon xanthomata. Of course, the added advantage of early cholesterol-lowering treatment is to decrease the chance of premature CHD.

Commonly used criteria for diagnosing FH include the Simon Broome Register criteria (Table 14.7) put forward in 1991 in the United Kingdom [89] and the Dutch Lipid Clinic Networks criteria for FH (Table 14.8) published in 1999 [90]. More recently alternative diagnostic criteria have been proposed [91].

**TABLE 14.7** Simon Broome familial hypercholesterolemia register diagnostic criteria for heterozygous familial hypercholesterolemia (HeFH).

Criteria	Description
A	Total cholesterol concentration >7.5 mmol/L in adults or >6.7 mmol/L in children aged <16 years, or Low-density lipoprotein cholesterol concentration >4.9 mmol/L in adults or >4.0 mmol/L in children
B	Tendinous xanthomata in patient or first-degree relative
C	DNA-based evidence of mutation in <i>LDLR</i>
D	Family history of myocardial infarction <50 years in second-degree relative or <60 years in first-degree relative
E	Family history of raised total cholesterol concentration $\geq 7.5$ mmol/L in first- or second-degree relative
<b>Diagnosis</b>	
A + B or C constitutes a definite diagnosis of HeFH	
A + D or A + E constitute probable HeFH	

With permission from Bhatnagar D, Soran H, Durrington PN. Hypercholesterolaemia and its management. *BMJ* 2008;337:a993.

**TABLE 14.8** Criteria for the diagnosis of heterozygous familial hypercholesterolemia according to Dutch Lipid Clinic Network.

Variable	Criteria	Score
Family history	First-degree relative with premature CAD*, or first degree relative with LDL-C >95th centile	1
	First-degree relative with tendon xanthomata and/or corneal arcus, or children <18 years with LDL-C >95th centile	2
Clinical history	Premature CAD*	2
	Premature cerebral or peripheral vascular disease*	1
Physical examination	Tendon xanthomata	6
	Corneal arcus <45 years	4
LDL-C	>8.5 mmol/L (>325 mg/dL)	8
	6.5–8.4 mmol/L (251–325 mg/dL)	5
	5.0–6.4 mmol/L (191–250 mg/dL)	3
	4.0–4.9 mmol/L (155–190 mg/dL)	1
DNA analysis	Functional mutation in <i>LDLR</i> , <i>apoB</i> , or <i>PCSK9</i> gene	8
Definite FH	Score >8	
Probable FH	Score 6–8	
Possible FH	Score 3–5	
No diagnosis	Score <3	

\* Premature = < 55 years in men; < 60 years in women.

CAD, Coronary artery disease; FH, familial hypercholesterolemia; FH, familial hypercholesterolemia; LDL-C, low-density lipoprotein cholesterol; LDLR, low-density lipoprotein receptor.

Adapted from 2016 ESC/EAS guidelines for the management of dyslipidemias. From Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, et al. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: Consensus Statement of the European Atherosclerosis Society. *Eur Heart J* 2013;34(45):3478–90. doi:10.1093/eurheartj/ehz273. Epub 2013 Aug 15.



## 14.9 Screening strategies for familial hypercholesterolemia

FH is a relatively common genetic disorder with prevalence between 1 in 200 and 1 in 500, but the population at risk is much smaller compared with other lipid-related coronary risk factors [92]. However, the relative odds of a patient with FH developing CHD are the highest compared with common hypercholesterolemia, as the exposure to elevated serum cholesterol is present from birth. Patients with FH, therefore, should be actively identified through a dedicated screening program, but a lack of resources has made it less of a priority for public health bodies. A wide variety of mutations causing FH and phenotypic variability have added to the difficulty in establishing universal diagnostic criteria. In routine clinical practice, patients with FH are mostly identified incidentally when presenting with premature CHD or are screened when a relative is found to have a marked raised serum cholesterol or CHD [93].

Screening relatives of FH probands or cascade testing or family tracing based on testing serum cholesterol concentrations or by genetic testing has been tried in several settings in the United Kingdom, Holland, Norway, Iceland, and the United States [93–95]. It is best carried out in specialist centers or clinics run by nurses trained in genetic counseling and cardiovascular risk management with access to genetic testing. Allelic heterogeneity and the other genetic and environmental influences may limit the utility of genetic testing and may be best applied to relatives of probands or in relatively homogenous communities. A detailed analysis of cost-effectiveness of screening for FH concludes that cascade testing using clinical criteria was the most cost-effective, followed closely by genetic cascade testing [96].

## 14.10 Genetic disorders resulting in hypertriglyceridemia

These rare disorders result from overproduction or disordered catabolism of the triglyceride-rich lipoproteins and their remnants resulting in moderate-to-severe hypertriglyceridemia (Table 14.9). Most have characteristic clinical features. The main risk from the moderate-to-severe hypertriglyceridemia is acute and recurrent pancreatitis. Establishing the etiology of these uncommon disorders using specialist biochemical analysis is difficult, and genetic testing is now the preferred method. Monogenic disorders are very rare. It has recently been suggested that almost all of those with severe hypertriglyceridemia who had genetically identified factors had polygenic hypertriglyceridemia with accumulation of common variants being the most predominant genetic determinant and that homozygous or compound heterozygous variants were infrequent [97]. Many patients with severe hypertriglyceridemia are identified in the clinical laboratory where the serum or plasma is markedly lipemic. Although genetic testing in specialist laboratories to find the cause of hypertriglyceridemia is now usual, in the past (and in some cases still), keeping the lipemic serum or plasma in the fridge could point to the disorder.

**TABLE 14.9** Summary of some primary lipoprotein disorders leading to hypertriglyceridemia.

Lipid disorder	Molecular defect	Incidence	Lipoprotein abnormality	Presentation
Type 3 hyperlipoproteinemia	Apo E2 homozygosity	1/5000	Increase in IDL and chylomicron remnants	Palmer crease xanthomata and tuberoeruptive xanthomata
Familial combined dyslipidemia	Unclear	1/200	Overproduction of VLDL and LDL	Coexists with obesity, impaired glucose tolerance, hypertension, and/or gout; leads to premature CHD
Familial hypertriglyceridemia	Unclear	1/500	Increased VLDL	Family members usually affected
FCS	Deficiency of lipoprotein lipase or its activator apo CII or GPIHBP1, APOA5, and LMF1	1/1,000,000	Increased chylomicrons	Marked hypertriglyceridemia, eruptive xanthomata, recurrent pancreatitis

FCS, Familial chylomicronemia syndrome; IDL, intermediate-density lipoproteins; LDL, low-density lipoprotein; VLDL, very low-density lipoprotein.

Adapted from Shah AS, Wilson DP. Primary hypertriglyceridemia in children and adolescents. *J Clin Lipidol* 2015;9(5 Suppl):S20–8. doi:10.1016/j.jacl.2015.04.004.

Where chylomicrons were increased, they would float to the top like cream, leaving the serum/plasma clear. If the hypertriglyceridemia was due to an increase in VLDL, this would generally result in uniform distribution of lipemia with no creamy layer on top.

### 14.11 Type 3 hyperlipoproteinemia

Type 3 hyperlipoproteinemia presents clinically as a mixed hyperlipidemia with rather high serum cholesterol and triglycerides. Its untreated form is generally characterized by eruptive xanthomata, palmar crease xanthomata, and tuberose xanthomata on elbows and knees [1]. It is a rare disorder with a prevalence of about 1 in 10,000. Type 3 hyperlipoproteinemia results from mutations in apolipoprotein E, which acts as a ligand for hepatic receptors for triglyceride-rich lipoproteins and their remnants (Fig. 14.1 and Table 14.1). The gene for apo E has three variants termed E2, E3, and E4 producing six possible genotypes E2//E2, E2/E3, E2/E4, E3/E3, E3/E4, E4/E4. The wild-type in humans is E3/E3. The E2 variant results in higher triglycerides and the E4 variant leads to higher serum cholesterol. The exact cause of type 3 hyperlipoproteinemia is not known, but nearly 10%–15% of those with an E2/E2 genotype develop the mixed hyperlipidemia, which is predominantly due to decreased clearance of remnant lipoproteins [98]. Impaired glucose tolerance, type 2 diabetes, and hypothyroidism appear to be common associations and may be triggers for the hyperlipidemia [1]. It is important to rule out type 3 hyperlipoproteinemia in patients with mixed hyperlipidemia as the condition leads to premature CHD and peripheral vascular disease, and patients respond well to treatment with fibrate and statin drugs.

The demonstration of cholesterol-enriched VLDL by ultracentrifugation is a “definitive” test but is not suitable for routine laboratory use. In the past detecting the apo E phenotype and the presence of a characteristic broad beta band on polyacrylamide gradient gel electrophoresis were used as diagnostic tests. Nowadays, a combination of apo E genotyping and examining the ratio of apo B to total cholesterol and/or triglycerides in serum can help to establish the diagnosis [99]. At times, the condition can be mistaken for heterozygous FH due to the presence of xanthomata or with FCH, which also presents with a mixed hyperlipidemia. In these hypercholesterolemia conditions the associated hypertriglyceridemia is generally not as prominent as seen in type 3 hyperlipoproteinemia.

Table 14.10 lists the other rare primary disorders of lipoprotein metabolism [100].

### 14.12 Treatment of primary lipoprotein disorders

Patients with primary disorders of lipoprotein metabolism are at considerable risk of developing premature CHD [101], pancreatitis, and rarely conditions, such as kidney disease or neuropathy. Most have what is classified as severe hyperlipidemia and lifestyle measures alone to treat elevated serum lipids are not sufficient. They will need high-intensity lipid-lowering drugs for effective lipid lowering and in many cases will end up on a combination of drugs. These patients are best managed in specialist centers.

### 14.13 Management of hypercholesterolemia

Table 14.11 lists the commonly used cholesterol-lowering drugs [102]. The mainstay of cholesterol lowering is the drugs known as *statins*. They decrease cholesterol synthesis in the liver by competitively inhibiting 3-hydroxy-methylglutaryl coenzyme A reductase, thereby depleting hepatic cholesterol. This results in upregulation of hepatic LDLRs expression leading to clearance of LDL from blood. Statins vary in their cholesterol-lowering potency [103,104]. They substantially reduce cardiovascular morbidity and mortality in both primary and secondary prevention settings. Several studies have shown conclusively that this mainly due to lowering of LDL-C [17–19,21]. In general, statins are safe and well tolerated. About 15%–20% of those receiving statins may have statin-associated muscle symptoms/statin intolerance [105,106]. The risk of myositis is lower with lower doses and does not relate to the degree of LDL-C lowering. Generally the excess risk of myopathy compared with placebo in statin randomized controlled trials (RCTs) is less than 1 in 10,000 patients treated with statins [107]. Muscle-aches and minor elevations of creatine kinase are common regardless of statin treatment but are more frequently attributed to statins in clinical practice than is the case in placebo-controlled trials.

**TABLE 14.10** Genetic, clinical, and biochemical features of rare genetic lipid disorders.

Disorder	Gene	Inheritance	Clinical phenotype	Biochemical phenotype				
				Features	TC	HDL-C	LDL-C	TG
Abetalipoproteinemia	MTTP	Recessive	Failure to thrive, nonspecific GI symptoms, steatorrhea, progressive neurological and ophthalmic abnormalities, easy bruising and bleeding, osteopenia	Absence of LDL-C, VLDL-C, chylomicrons, and apoB, low total cholesterol and triglycerides, very low to absent vitamins A, D, and E, prolonged INR	↓↓↓	↓	↓↓↓	↓↓↓
Homozygous familial hypobetalipoproteinemia	APOB	Codominant	As for abetalipoproteinemia	As for abetalipoproteinemia	↓↓↓	↓	↓↓↓	↓↓↓
Heterozygous familial hypobetalipoproteinemia	APOB	Codominant	Asymptomatic; hepatic steatosis	Less than one-third normal LDL-C, low vitamin E	↓	n	↓↓	↓
Chylomicron retention disease	SARIB	Recessive	Failure to thrive, nonspecific GI symptoms, steatorrhea, hepatomegaly, neurological, muscular, and ophthalmic complications	Absence of chylomicrons, LDL-C <50% levels of normal for age and sex	↓↓	↓↓	↓↓	n
Familial chylomicronemia syndrome	LPL APOC2 APOA5 GPIHBP1 LMFI	Recessive	Failure to thrive, nonspecific GI symptoms, hepatosplenomegaly, lipemia retinalis, eruptive xanthomas, pancreatitis	Severe hypertriglyceridemia (> 10 mmol/L); fasting chylomicronemia	↑↑	↓↓	n-↓	↑↑↑
Hepatic lipase deficiency	LIPC	Recessive	Asymptomatic, xanthomas, atherosclerosis	High HDL-C, hypertriglyceridemia, abnormally large TG-rich HDL and LDL particles	↑	↑	n-↓	↑
Tangier disease	ABCA1	Recessive	Hyperplastic orange tonsils, peripheral neuropathy, hepatosplenomegaly, rectal mucosa changes, corneal opacities, hematological abnormalities, premature CAD	Very reduced HDL-C and apoA-I, hypertriglyceridemia	↓	↓↓↓	n-↓	↑
Familial apoA-I deficiency	APOA1	Codominant	Premature CAD, xanthomas (cutaneous, tendinous, tuberoeruptive), retinal lipid deposition, corneal opacification, neurological abnormalities, amyloidogenesis	Very low (homozygotes) or half-normal HDL-C (heterozygotes), low apoA-I	↑	↓↓	n-↓	n
Familial LCAT deficiency (complete)	LCAT	Recessive	Corneal opacification, kidney disease, hemolytic anemia, splenomegaly	Very low HDL-C and apoA-I, increased plasma FC/CE ratio, very low cholesterol esterification rate, presence of LpX	↑	↓↓	n-↓	n-↑
Fish eye disease	LCAT	Recessive	Severe corneal opacification, premature CAD	Very low HDL-C and apoA-I levels	↑	↓↓	n-↓	n-↑
Homozygous familial hypercholesterolemia	LDLR PCSK9 APOB	Codominant	Extensive xanthomas, arcus cornealis, premature CAD, valvular heart disease	Extreme LDL-C elevation	↑↑↑	↓	↑↑↑	n
Autosomal recessive hypercholesterolemia	LDLRAP1	Recessive	Extensive xanthomas, arcus cornealis, premature CAD, valvular heart disease	Extreme LDL-C elevation	↑↑↑	↓-n	↑↑↑	n
Sitosterolemia	ABCG5 ABCG8	Recessive	Xanthomas, arthralgia, hemolytic anemia, macrothrombocytopenia, splenomegaly, premature CAD	High plant sterols, hypercholesterolemia	↓ (infant ↑↑↑)	↓-n	↓ (infant ↑↑↑)	n
Familial dysbetalipoproteinemia	APOE	Other	Requires trigger by secondary cause; xanthomas (palmar crease, tuberous, eruptive, tendinous), hepatosplenomegaly, premature CAD	Raised TC and TG (approx. 2:1 ratio), significantly increased IDL (↑↑↑)	↑↑	↓	↓ (IDL) ↑↑↑	↑↑
Lysosomal acid lipase deficiency	LIPA	Recessive	Severity varies; failure to thrive, hepatosplenomegaly, steatorrhea, thrombocytopenia, liver disease, nonspecific GI symptoms, premature CAD	Reduced lysosomal acid lipase activity, raised TC and TG, elevated serum transaminases	n-↑	↓	↑	↑

↓, Decreased; ↑, increased; n, normal. *CAD*, Coronary artery disease; *FC/CE*, free cholesterol/cholesterol ester; *HDL-C*, high-density lipoprotein-cholesterol; *IDL*, intermediate-density lipoprotein; *INR*, international normalized ratio; *LCAT*, lecithin:cholesterol acyltransferase; *LDL-C*, low-density lipoprotein-cholesterol; *TC*, total cholesterol; *TG*, triglycerides; *VLDL-C*, very low-density lipoprotein cholesterol.

With permission from Ng DM, Burnett JR, Bell DA, Hegele RA, Hooper AJ. Update on the diagnosis, treatment and management of rare genetic lipid disorders. *Pathology* 2019;51(2):193–201. Available from: <https://doi.org/10.1016/j.pathol.2018.11.005>.

**TABLE 14.11** Characteristics of common lipid-lowering medications that are used to lower low-density lipoprotein cholesterol (LDL-C).<sup>a</sup>

Medication class	Mechanism of action	Drugs	Total daily dose range (mg/d) <sup>b</sup>	Dosing frequency	Comments
HMG-CoA reductase inhibitors (also known as statins)	Competitively inhibit HMG-CoA reductase (rate-limiting step of endogenous cholesterol production); increase the number of LDL receptors	Atorvastatin	10–80	Once daily	<ul style="list-style-type: none"> <li>First-line therapy for nearly all patients, as based on extensive evidence demonstrating reductions in cardiovascular events over wide range of LDL-C and overall safety</li> <li>Potential LDL-C reduction<sup>c</sup> is 18%–55%</li> <li>LDL-C reductions vary according to dose of the specific statin</li> <li>Fluvastatin, lovastatin, pravastatin, and simvastatin have short half-lives. They should be administered in the evening to achieve maximum LDL-C reduction. Atorvastatin, fluvastatin XL, pitavastatin, and rosuvastatin can be dosed anytime of the day</li> </ul>
		Fluvastatin	20–80	Once or twice daily	
		Lovastatin	10–80	Once or twice daily	
		Pitavastatin	1–4	Once daily	
		Pravastatin	10–80	Once daily	
		Rosuvastatin	5–40	Once daily	
		Simvastatin	5–40	Once daily	
Bile acid sequestrants	Bind bile acids in the gut, interrupt enterohepatic recirculation of bile acids and impede their reabsorption, decrease bile acid pooling in the liver, increase conversion of cholesterol to bile acids, increase the number of LDL receptors	Cholestyramine	4000–24,000	Once or twice daily	<ul style="list-style-type: none"> <li>Nonsystemic add-ons to statin therapy, or used in patients with statin-associated side effects, including statin-associated muscle symptoms</li> <li>Potential LDL-C reduction<sup>c</sup> is 15%–30%</li> <li>Available as tablets or powder for suspension</li> <li>Gastrointestinal side effects may limit use</li> <li>May increase serum TG levels; avoid if TG &gt;300 mg/dL</li> <li>Colesevelam is approved for use in type 2 diabetes mellitus to reduce hemoglobin A1C</li> <li>Can bind absorption of other medications (less with colesevelam); should be administered at least 1 h before or 4 h after other medications to minimize potential drug–drug interaction</li> </ul>
		Colesevelam	3750	Once or twice daily	
		Colestipol	5000–30,000	Once to six times daily	
Cholesterol absorption inhibitors	Block the cholesterol transport Niemann–Pick C1-Like 1 protein to inhibit intestinal and biliary cholesterol absorption; increase the number of LDL receptors	Ezetimibe	10	Once daily	<ul style="list-style-type: none"> <li>Evidence-based add-on to statin therapy in very high-risk patients or in patients with statin-associated side effects, including statin-associated muscle symptoms</li> <li>Potential LDL-C reduction<sup>c</sup> is 13%–20%</li> <li>Approved for use in homozygous sitosterolemia to reduce elevated sitosterol and campesterol</li> </ul>

(Continued)

TABLE 14.11 (Continued)

Medication class	Mechanism of action	Drugs	Total daily dose range (mg/d) <sup>b</sup>	Dosing frequency	Comments
PCSK9 inhibitors	Fully human monoclonal antibodies that bind to PCSK9 and decrease degradation of the LDL receptor	Alirocumab	75–150	Every 2 weeks	<ul style="list-style-type: none"> <li>Evidence-based add-on to statin therapy in very high-risk patients</li> <li>Potential LDL-C reduction<sup>c</sup> is 43%–64%</li> <li>Lower LDL-C reduction in heterozygous FH when added to tolerated statin/ezetimibe therapy</li> <li>Mean LDL-C reduction is 30% with evolocumab in homozygous FH (50)</li> <li>Requires subcutaneous injection</li> </ul>
			300	Every 4 weeks	
		Evolocumab	140	Every 2 weeks	
			420	Every 4 weeks	

<sup>a</sup>Lomitapide and mipomersen sodium are other medications that are used to lower LDL-C in patients with homozygous familial hypercholesterolemia. Though rarely prescribed, these medications are usually prescribed by a clinical lipidologist because of restricted access through a Risk Evaluation and Mitigation Strategy program to assure safe use.

<sup>b</sup>Dosages and administration from FDA-approved labeling [available from: <http://dailymed.nlm.nih.gov/dailymed/index.cfm> (51)].

<sup>c</sup>Potential LDL-C lowering based on estimations from the National Lipid Association (52), or product labeling.

FDA, US Food and Drug Administration; FH, familial hypercholesterolemia; HMG-CoA, 3-hydroxy-3-methylglutaryl-coenzyme; PCSK9, proprotein convertase subtilisin/kexin type 9; TG, triglycerides; XL, extended release.

With permission from the web supplement to Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2018. pii: S0735-1097(18)39033-39038. Available from: <https://doi.org/10.1016/j.jacc.2018.11.002> [accessed 09.06.19].

Ezetimibe lowers serum cholesterol by inhibiting intestinal dietary and biliary cholesterol absorption by binding to the Niemann–Pick C1-like 1 receptors. Its lowers LDL-C-lowering agent by about 10%–15% [108] and is best used in combination with statin therapy in patients with particularly high LDL-C levels and in high-risk patients who are truly statin intolerant [36,109]. There is evidence of cardiovascular event reduction in clinical trials when used alongside simvastatin [110,111].

Bile acid–sequestering agents are ion exchange resins that prevent the reabsorption of bile acids from the terminal ileum. This leads to increased hepatic requirement for cholesterol for bile acid synthesis to replenish the enterohepatic pool and an increase in LDLR expression causing lowering of circulating LDL particles with its cargo [112]. Before the advent of statins, they were the mainstay of for cholesterol-lowering, but they are poorly tolerated due to gastrointestinal side effects.

PCSK9 inhibitors [113] have recently been licensed in the United Kingdom, but they are only to be used in patients with extremely high CVD risk as specified by NICE [114,115].

There has been recent interest in a possible role of *nutraceuticals and functional foods* as an adjunct to lipid-lowering therapy, though their efficacy and safety remain poorly understood [116].

Surgery, liver transplantation, and extracorporeal LDL removal have all been used to treat severe hypercholesterolemia. Ileal bypass, a procedure now rarely considered, effectively reduces LDL-C and lowers CHD mortality by 35% [117]. Liver transplantation has been used in homozygous FH patients to provide the functional hepatic LDLRs that these patients lack [118]. Bariatric surgery has also been shown to be effective at reducing LDL-C in obese patients. A recent meta-analysis (48 studies, 6077 participants) showed a significant reduction in LDL-C by 1-month postoperatively (standardized mean difference −0.92, 95% CI: −1.31 to −0.52) with the effect being maintained in the longer term [119]. Indeed, in addition to improvements in the lipid profile, bariatric surgery–induced weight loss is also associated with significant improvements in systemic inflammation, insulin resistance, mediators of vascular inflammation, vascular function, perivascular adipose tissue inflammation, and

adipose tissue anticontractile properties [120,121]. Where available, for patients with homozygous FH and severe heterozygous FH, *lipoprotein apheresis* is a well-established procedure that can decrease mean LDL-C by more than 50% and reduce mortality [122,123]. Access to this procedure remains limited because of its relatively high cost and low availability. *Lomitapide* is a small-molecule inhibitor of microsomal triglyceride transfer protein that reduces the hepatic assembly of VLDL and intestinal chylomicrons and consequently reduces LDL-C production, an action independent of LDLR activity. *Lomitapide* is a high-cost drug licensed for patients with homozygous FH [124,125]. *Mipomersen*, a second-generation antisense oligonucleotide, which inhibits hepatic APOB synthesis, has a similar effect [126] but is not licensed in Europe.

#### 14.14 Agents in development

Infusion of HDL mimetic peptides is also currently being investigated [127–131]. Improving the functionality of HDL may be more important than increasing HDL-C [5]. Gene therapy, supplying the normal LDLR gene via a plasmid, is also a potential future treatment [132]. In addition, RVX-208, which has been shown to improve the particle profile of HDL to favor one which promotes RCT, has shown encouraging early results in potentially preventing and treating atherosclerosis [133,134]. In a phase 2 trial a synthetic small interfering RNA directed against PCSK9 was shown both to reduce PCSK9 and LDL-C levels [135]. In addition, another novel agent, bempedoic acid, inhibits ATP citrate lyase, an enzyme involved in fatty acid and cholesterol synthesis, predominantly in the liver and white adipose tissue and has been shown to reduce LDL-C in several phase 2 trials [136]. The CETP inhibitor Anacetrapib may reduce CVD risk as reported in the REVEAL trial [137].

#### 14.15 Management of hypertriglyceridemia

*Fibrates* (*fenofibrate*, *bezafibrate*, *gemfibrozil*) and *nicotinic acid* (and its derivative *acipimox*) are effective triglyceride-lowering drugs, but clinical trial evidence does not support their use to lower CVD risk except in a subset of patients [138–141]. Purified *omega-3 fatty acids* can decrease triglycerides [142] and decreased CVD risk in some studies [143]. However, many patients with primary disorders of triglyceride metabolism are at risk of pancreatitis, and in the presence of hypercholesterolemia, mild-to-moderate increases in serum triglycerides increase coronary risk. Triglyceride lowering is, therefore, best used in patients with mixed hyperlipidemia and in those who have had an episode of pancreatitis or are prone to recurrent pancreatitis. The mixed hyperlipidemia seen in patients with type 3 hyperlipoproteinemia responds sharply to fibrate drugs. hypertriglyceridemia is pronounced postprandially, particularly after large intake of fats or alcohol, so patients are advised to abstain from alcohol and follow a low-fat diet for the triglyceride-lowering drugs to work effectively [144]. The potent statins such as atorvastatin and rosuvastatin have useful triglyceride-lowering properties. Their use will result in lowering LDL-C and, therefore, coronary risk, but the persistent increase in serum triglycerides may need the addition of fibrates. This combination of fibrates with statins has the potential of severe muscle toxicity (particularly with gemfibrozil, which should *not* be used with statin drugs) and should be used with caution. Purified omega-3 fatty acids, such as Omacor [145] or icosapent ethyl [146], can be combined with statins in mixed hyperlipidemia where the statin-fibrate is not appropriate. Nicotinic acid can also be combined with statins, but side effects related to nicotinic acid remain problematic.

In patients with severe hypertriglyceridemia, particularly associated with frequent bouts of pancreatitis, plasma exchange may be of use, although most patients also respond to a regime of nil by mouth and intravenous fluids.

Several new drugs such as inhibitors of apo C3 (volanesorsen) and ANGPTL3 (Evinacumab) are in development, but they are still experimental [147].

## References

- [1] Durrington PN. Diagnosis and management. Hyperlipidaemia. 3rd ed. London: Hodder Arnold; 2007.
- [2] Bhatnagar D. Lipid-lowering drugs in the management of hyperlipidaemia. *Pharmacol Ther* 1998;79:205–30.
- [3] Langlois MR, Chapman MJ, Cobbaert C, Mora S, Remaley AT, Ros E, et al. Quantifying atherogenic lipoproteins: current and future challenges in the era of personalized medicine and very low concentrations of LDL cholesterol. A consensus statement from EAS and EFLM. *Clin Chem* 2018;64(7):1006–33. Available from: <https://doi.org/10.1373/clinchem.2018.287037>.



- [4] Charlton-Menys V, Durrington PN. Human cholesterol metabolism and therapeutic molecules. *Exp Physiol* 2008;93(1):27–42. Available from: <https://doi.org/10.1113/expphysiol.2007.035147>.
- [5] Soran H, Hama S, Yadav R, Durrington PN. HDL functionality. *Curr Opin Lipidol* 2012;23(4):353–66. Available from: <https://doi.org/10.1097/MOL.0b013e328355ca25>.
- [6] Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. *Br Med J* 1991;303:276–82.
- [7] Iso H. Changes in coronary heart disease risk among Japanese. *Circulation* 2008;118:2725–9.
- [8] Li X, Mai J, Rao X, et al. Adverse trends of cardiovascular risk factors among low risk populations (1983–1994): a cohort study of workers and farmers in Guangzhou, China. *BMC Public Health* 2011;11:931.
- [9] BHF. BHF CVD statistics compendium 2017. BHF; 2017. [Internet]. Available from: <https://www.bhf.org.uk/research/heart-statistics/heart-statistics-publications/cardiovascular-disease-statistics-2017>.
- [10] Narayan KM, Ali MK, Koplan JP. Global noncommunicable diseases—where worlds meet. *N Engl J Med* 2010;363:1196–8.
- [11] Reddy KS. Cardiovascular diseases in the developing countries: dimensions, determinants, dynamics and directions for public health action. *Public Health Nutr* 2002;5:231–7.
- [12] Mendis S. Cardiovascular risk assessment and management in developing countries. *Vasc Health Risk Manage* 2005;1:15–18.
- [13] Robertson TL, Kato H, Gordon T, et al. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California. Coronary heart disease risk factors in Japan and Hawaii. *Am J Cardiol* 1977;39:244–9.
- [14] Bhatnagar D, Anand IS, Durrington PN, Patel DJ, Wander GS, Mackness MI, et al. Coronary risk factors in people from the Indian sub-continent living in west London and their siblings in India. *Lancet* 1995;345(8947):405–9.
- [15] Patel JV, Vyas A, Cruickshank JK, et al. Impact of migration on coronary heart disease risk factors: comparison of Gujaratis in Britain and their contemporaries in villages of origin in India. *Atherosclerosis* 2006;185:297–306.
- [16] Bhatnagar D, Soran H, Durrington PN. Hypercholesterolaemia and its management. *Br Med J* 2008;337:a993.
- [17] Cholesterol Treatment Trialists Collaborators, Kearney PM, Blackwell L, Collins R, et al. Efficacy of cholesterol-lowering therapy in 18,686 people with diabetes in 14 randomised trials of statins: a meta-analysis. *Lancet* 2008;371:117–25.
- [18] Cholesterol Treatment Trialists Collaborators, Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* 2010;376:1670–81.
- [19] Colantonio LD, Bittner V, Reynolds K, Levitan EB, Rosenson RS, Banach M, et al. Association of serum lipids and coronary heart disease in contemporary observational studies. *Circulation* 2016;133:256–64. Available from: <https://doi.org/10.1161/CIRCULATIONAHA.115.011646>.
- [20] Soran H, Kwok S, Adam S, Ho JH, Durrington PN. Evidence for more intensive cholesterol lowering. *Curr Opin Lipidol* 2017;28:291–9.
- [21] Soran H, Dent R, Durrington P. Evidence-based goals in LDL-C reduction. *Clin Res Cardiol* 2017;106:237–48.
- [22] Stamler J, Wentworth D, Neaton JD. Is the relationship between serum cholesterol and risk of premature death from coronary heart disease continuous and graded? Findings in 356,222 primary screenees of the multiple risk factor intervention trial (MRFIT). *J Am Med Assoc* 1986;256:2823–8.
- [23] LaRosa JC, Hunninghake D, Bush D, Criqui MH, Getz GS, Gotto Jr AM, et al. The cholesterol facts. A summary of the evidence relating dietary fats, serum cholesterol, and coronary heart disease. A joint statement by the American Heart Association and the National Heart, Lung, and Blood Institute. The Task Force on Cholesterol Issues, American Heart Association. *Circulation* 1990;81:1721–33.
- [24] Jeemon P, Neogi S, Bhatnagar D, Cruickshank JK, Prabhakaran D. The impact of migration on cardiovascular diseases and its risk factors among people of Indian origin. *Curr Sci* 2009;97:378–84.
- [25] Elder SJ, Lichtenstein AH, Pittas AG, et al. Genetic and environmental influences on factors associated with cardiovascular disease and the metabolic syndrome. *J Lipid Res* 2009;50:1917–26.
- [26] Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science* 1986;232:34–47.
- [27] Hegele RA. Plasma lipoproteins: genetic influences and clinical implications. *Nat Rev Genet* 2009;10(2):109–21. Available from: <https://doi.org/10.1038/nrg2481>.
- [28] Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet* 2018;50(11):1514–23. Available from: <https://doi.org/10.1038/s41588-018-0222-9>.
- [29] Dron JS, Hegel RA. Genetics of lipid and lipoprotein disorders and traits. *Curr Genet Med Rep* 2016;4(3):130–41. Available from: <https://doi.org/10.1007/s40142-016-0097>.
- [30] Dron JS, Hegel RA. Polygenic influences on dyslipidemias. *Curr Opin Lipidol* 2018;29(2):133–43. Available from: <https://doi.org/10.1097/MOL.0000000000000482>.
- [31] Dron JS, Hegel RA. The evolution of genetic-based risk scores for lipids and cardiovascular disease. *Curr Opin Lipidol* 2019;30:71–81. Available from: <https://doi.org/10.1097/MOL.0000000000000576>.
- [32] National Institute for Health and Clinical Excellence (NICE). Cardiovascular disease: risk assessment and reduction, including lipid modification. Clinical guideline [CG181]. NICE: London; 2014. Available from: <https://www.nice.org.uk/guidance/cg181> [updated Sep 2016].
- [33] Catapano AL, Graham I, De Backer G, et al. 2016 ESC/EAS Guidelines for the management of dyslipidaemias. *Eur Heart J* 2016;37:2999–3058.
- [34] Board JBS. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). *Heart* 2014;100(Suppl. 2):ii1–67.
- [35] Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. *Br Med J* 2012;344:e4181.
- [36] Authors/Task Force M, Catapano AL, Graham I, De Backer G, et al. 2016 ESC/EAS guidelines for the management of dyslipidaemias: the Task Force for the Management of Dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Atherosclerosis* 2016;253:281–344.

- [37] Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014;129(25 Suppl. 2):S1–45.
- [38] Goff Jr DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014;129(25 Suppl. 2):S49–73.
- [39] Singh N, Kwok S, Seneviratne CJ, France M, Durrington P. The National Cholesterol Education Program III scoring system for CHD risk estimation cannot be used with European recommendations. *Br J Cardiol* 2004;11:282–6.
- [40] Soran H, Schofield JD, Durrington PN. Cholesterol, not just cardiovascular risk, is important in deciding who should receive statin treatment. *Eur Heart J* 2015;36(43):2975–83. Available from: <https://doi.org/10.1093/eurheartj/ehv340> Epub 2015 Aug 4.
- [41] Soran H, Adam S, Durrington PN. Are recent statin recommendations to employ fixed doses and abandon targets effective for treatment of hypercholesterolaemia? Investigation based on number needed to treat. *Eur J Prev Cardiol* 2017;24(1):76–83 Epub 2016 Sep 8.
- [42] Soran H, Adam S, Durrington PN. Optimising treatment of hyperlipidaemia: quantitative evaluation of UK, USA and European guidelines taking account of both LDL cholesterol levels and cardiovascular disease risk. *Atherosclerosis* 2018;278:135–42. Available from: <https://doi.org/10.1016/j.atherosclerosis.2018.08.040> Epub 2018 Aug.
- [43] Asselbergs FW, Guo Y, van Iperen EP, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am J Hum Genet* 2012;91:823–38 [PMC free article] [PubMed] [Google Scholar].
- [44] Talmud PJ, Shah S, Whittall R, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. *Lancet* 2013;381:1293–301.
- [45] Miller M, Stone NJ, Ballantyne C, et al. Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* 2011;123:2292–333.
- [46] Sandesara PB, Virani SS, Fazio S, Shapiro MD. The forgotten lipids: triglycerides, remnant cholesterol, an atherosclerotic cardiovascular disease risk. *Endocrinol Rev* 2019;40(2):537–57. Available from: <https://doi.org/10.1210/er.2018-00184>.
- [47] Nordestgaard BG. Triglyceride-rich lipoproteins an atherosclerotic cardiovascular disease: new insights from epidemiology, genetics, and biology. *Circ Res* 2016;118(4):547–63. Available from: <https://doi.org/10.1161/CIRCRESAHA.115.306249>.
- [48] Goldberg IJ. 2017 Georg Lyman duff memorial lecture: fat in the blood, fat in the artery, fat in the heart: triglyceride in physiology and disease. *Arterioscler Thromb Vasc Biol* 2018;38(4):700–6. Available from: <https://doi.org/10.1161/ATVBAHA.117.309666>.
- [49] Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009;120:1640–5.
- [50] Huang PL. A comprehensive definition for metabolic syndrome. *Dis Model Mech* 2009;2:231–7.
- [51] Ford ES, Li C, Sattar N. Metabolic syndrome and incident diabetes: current state of the evidence. *Diabetes Care* 2008;31:1898–904.
- [52] Soran H, Schofield JD, Adam S, Durrington PN. Diabetic dyslipidaemia. *Curr Opin Lipidol* 2016;27:313–22 [PubMed] [Google Scholar].
- [53] Schofield JD, Liu Y, Rao-Balakrishna P, Malik RA, Soran H. Diabetes dyslipidemia. *Diabetes Ther* 2016;7:203–19 [PMC free article] [PubMed] [Google Scholar].
- [54] Brahm AJ, Hegele RA. Combined hyperlipidemia: familial but not (usually) monogenic. *Curr Opin Lipidol* 2016;27:131–40.
- [55] Goldstein JL, Hobbs HH, Brown MS. Familial hypercholesterolaemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. *The metabolic and molecular bases of inherited disease*. 8th ed. New York: McGraw-Hill; 2001. p. 2863–913.
- [56] Austin MA, Hutter CM, Zimmern RL, Humphries SE. Familial hypercholesterolemia and coronary heart disease: a HuGE association review. *Am J Epidemiol* 2004;160:421–9.
- [57] Lindgren V, Luskey KL, Russell DW, Francke U. Human genes involved in cholesterol metabolism: chromosomal mapping of the loci for the low density lipoprotein receptor and 3-hydroxy-3-methylglutaryl-coenzyme A reductase with cDNA probes. *Proc Natl Acad Sci USA* 1985;82:8567–71.
- [58] Yamamoto T, Davis CG, Brown MS, et al. The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* 1984;39:27–38.
- [59] Schneider WJ, Beisiegel U, Goldstein JL, Brown MS. Purification of the low density lipoprotein receptor, an acidic glycoprotein of 164,000 molecular weight. *J Biol Chem* 1982;257:2664–73.
- [60] Austin MA, Hutter CM, Zimmern RL, Humphries SE. Genetic causes of monogenic heterozygous familial hypercholesterolemia: a HuGE prevalence review. *Am J Epidemiol* 2004;160:407–20.
- [61] Bhatnagar D. Diagnosis and screening for familial hypercholesterolaemia. *Ann Clin Biochem* 2006;43:441–56.
- [62] Langer T, Strober W, Levy RI. The metabolism of low density lipoprotein in familial type II hyperlipoproteinemia. *J Clin Invest* 1972;51:1528–36.
- [63] Soutar AK, Myant NB, Thompson GR. Simultaneous measurement of apolipoprotein B turnover in very-low and low-density lipoproteins in familial hypercholesterolaemia. *Atherosclerosis* 1977;28:247–56.
- [64] Utermann G, Hoppichler F, Dieplinger H, Seed M, Thompson G, Boerwinkle E. Defects in the low density lipoprotein receptor gene affect lipoprotein (a) levels: multiplicative interaction of two gene loci associated with premature atherosclerosis. *Proc Natl Acad Sci USA* 1989;86:4171–4.
- [65] Mbewu AD, Bhatnagar D, Durrington PN, et al. Serum lipoprotein (a) in patients heterozygous for familial hypercholesterolemia, their relatives, and unrelated control populations. *Arterioscler Thromb* 1991;11:940–6.
- [66] Thompson GR, Seed M, Nithyananthan S, McCarthy S, Thorogood M. Genotypic and phenotypic variation in familial hypercholesterolemia. *Arteriosclerosis* 1989;9(1 Suppl.):175–180.
- [67] Hill JS, Hayden MR, Frohlich J, Pritchard PH. Genetic and environmental factors affecting the incidence of coronary artery disease in heterozygous familial hypercholesterolemia. *Arterioscler Thromb* 1991;11:290–7.

- [68] Beppu S, Minura Y, Sakakibara H, et al. Supravalvular aortic stenosis and coronary ostial stenosis in familial hypercholesterolemia: two-dimensional echocardiographic assessment. *Circulation* 1983;67:878–84.
- [69] Goldstein JL. The cardiac manifestations of the homozygous and heterozygous forms of familial type II hypercholesterolaemia. *Birth Defects* 1972;8:202.
- [70] Beaumont V, Jacotot B, Beaumont JL. Ischaemic disease in men and women with familial hypercholesterolaemia and xanthomatosis. A comparative study of genetic and environmental factors in 274 heterozygous cases. *Atherosclerosis* 1976;24:441–50.
- [71] Mabuchi H, Miyamoto S, Ueda K, et al. Causes of death in patients with familial hypercholesterolemia. *Atherosclerosis* 1986;61:1–6.
- [72] Hutter CM, Austin MA, Humphries SE. Familial hypercholesterolemia, peripheral arterial disease, and stroke: a HuGE minireview. *Am J Epidemiol* 2004;160:430–5.
- [73] Koeijvoets KC, Wiegman A, Rodenburg J, Defesche JC, Kastelein JJ, Sijbrands EJ. Effect of low-density lipoprotein receptor mutation on lipoproteins and cardiovascular disease risk: a parent–off spring study. *Atherosclerosis* 2005;180:93–9.
- [74] Huxley RR, Hawkins MH, Humphries SE, Karpe F, Neil HA, Simon Broome Familial Hyperlipidaemia Register Group, et al. Risk of fatal stroke in patients with treated familial hypercholesterolemia: a prospective registry study. *Stroke* 2003;34:22–5 Erratum in: *Stroke* 34, 826.
- [75] Stone NJ, Levy RI, Fredrickson DS, Verter J. Coronary artery disease in 116 kindred with familial type II hyperlipoproteinemia. *Circulation* 1974;49:476–88.
- [76] Slack J. Risks of ischaemic heart-disease in familial hyperlipoproteinaemic states. *Lancet* 1969;2:1380–2.
- [77] Jensen J, Blankenhorn DH, Kornerup V. Coronary disease in familial hypercholesterolaemia. *Circulation* 1967;36:77–82.
- [78] Heiberg A. The risk of atherosclerotic vascular disease in subjects with xanthomatosis. *Acta Med Scand* 1975;198:249–61.
- [79] Gudnason V, Day IN, Humphries SE. Effect on plasma lipid levels of different classes of mutations in the low-density lipoprotein receptor gene in patients with familial hypercholesterolemia. *Arterioscler Thromb* 1994;14:1717–22.
- [80] Vuorio AF, Aalto-Setälä K, Koivisto UM, et al. Familial hypercholesterolaemia in Finland: common, rare and mild mutations of the LDL receptor and their clinical consequences. Finnish FH-group. *Ann Med* 2001;33:410–21.
- [81] Sun XM, Patel DD, Bhatnagar D, et al. Characterization of a splice-site mutation in the gene for the LDL receptor associated with an unpredictably severe clinical phenotype in English patients with heterozygous FH. *Arterioscler Thromb Vasc Biol* 1995;15:219–27.
- [82] Sprecher DL, Schaefer EJ, Kent KM, et al. Cardiovascular features of homozygous familial hypercholesterolemia: analysis of 16 patients. *Am J Cardiol* 1984;54:20–30.
- [83] Kotze MJ, De Villiers WJ, Steyn K, et al. Phenotypic variation among familial hypercholesterolemics heterozygous for either one of two Afrikaner founder LDL receptor mutations. *Arterioscler Thromb* 1993;13:1460–8.
- [84] Vohl MC, Gaudet D, Moorjani S, et al. Comparison of the effect of two low-density lipoprotein receptor class mutations on coronary heart disease among French–Canadian patients heterozygous for familial hypercholesterolaemia. *Eur J Clin Invest* 1997;27:366–73.
- [85] Gaudet D, Vohl MC, Couture P, et al. Contribution of receptor negative versus receptor defective mutations in the LDL-receptor gene to angiographically assessed coronary artery disease among young (25–49 years) versus middle-aged (50–64 years) men. *Atherosclerosis* 1999;143:153–61.
- [86] Bertolini S, Cantafora A, Averna M, et al. Clinical expression of familial hypercholesterolemia in clusters of mutations of the LDL receptor gene that cause a receptor-defective or receptor negative phenotype. *Arterioscler Thromb Vasc Biol* 2000;20:E41–52.
- [87] Pimstone SN, Sun XM, du Souich C, Frohlich JJ, Hayden MR, Soutar AK. Phenotypic variation in heterozygous familial hypercholesterolemia: a comparison of Chinese patients with the same or similar mutations in the LDL receptor gene in China or Canada. *Arterioscler Thromb Vasc Biol* 1998;18:309–15.
- [88] Pereira E, Ferreira R, Hermelin B, et al. Recurrent and novel LDL receptor gene mutations causing heterozygous familial hypercholesterolemia in La Habana. *Hum Genet* 1995;96:319–22.
- [89] Scientific Steering Committee on behalf of the Simon Broome Register Group. Risk of fatal coronary heart disease in familial hypercholesterolaemia. *Br Med J* 1991;303:893–6.
- [90] World Health Organization. Familial hypercholesterolaemia; report of a second WHO consultation, WHO/HGN/FH/CONS/99.2. Geneva: World Health Organization; 1999.
- [91] Wald DS, Bestwick JP, Wald NJ. Child-parent familial hypercholesterolemia screening in primary care. *N Engl J Med* 2017;376(5):499–500. Available from: <https://doi.org/10.1056/NEJMc1615365>.
- [92] Gidding SS, Champagne MA, de Ferranti SD, Defesche J, Ito MK, Knowles JW, et al. The agenda for familial hypercholesterolemia: a scientific statement from the American Heart Association. *Circulation* 2015;132(22):2167–92. Available from: <https://doi.org/10.1161/CIR.0000000000000297>.
- [93] Akioyamen LE, Genest J, Shan SD, Reel RL, Albaum JM, Chu A, et al. Estimating the prevalence of heterozygous familial hypercholesterolaemia: a systematic review and meta-analysis. *BMJ Open* 2017;7(9):e016461. Available from: <https://doi.org/10.1136/bmjopen-2017-016461>.
- [94] Bhatnagar D, Morgan J, Siddiq S, Mackness MI, Miller JP, Durrington PN. Outcome of case finding among relatives of patients with known heterozygous familial hypercholesterolaemia. *Br Med J* 2000;321:1497–500.
- [95] Leren TP. Cascade genetic screening for familial hypercholesterolemia. *Clin Genet* 2004;66:483–7.
- [96] Hadfield SG, Humphries SE. Implementation of cascade testing for the detection of familial hypercholesterolaemia. *Curr Opin Lipidol* 2005;16:428–33.
- [97] Marks D, Wonderling D, Thorogood M, et al. Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2000;4:1–123.
- [98] Dron JS, Wang J, Cao H, McIntyre AD, Iacocca MA, Menard JR, et al. Severe hypertriglyceridemia is primarily polygenic. *J Clin Lipidol* 2019;13(1):80–8. Available from: <https://doi.org/10.1016/j.jacl.2018.10.006.98>.
- [99] Koopal C, Marais AD, Visseren FL. Familial dysbetalipoproteinemia: an underdiagnosed lipid disorder. *Curr Opin Endocrinol Diabetes Obes* 2017;24(2):133–9. Available from: <https://doi.org/10.1097/MED.0000000000000316>.

- [99] Sniderman AD, de Graaf J, Thanassoulis G, Tremblay AJ, Martin SS, Couture P. The spectrum of type III hyperlipoproteinemia. *J Clin Lipidol* 2018;12(6):1383–9. Available from: <https://doi.org/10.1016/j.jacl.2018.09.006>.
- [100] Ng DM, Burnett JR, Bell DA, Hegele RA, Hooper AJ. Update on the diagnosis, treatment and management of rare genetic lipid disorders. *Pathology* 2019;51(2):193–201. Available from: <https://doi.org/10.1016/j.pathol.2018.11.005>.
- [101] Humphries SE, Cooper JA, Capps N, Durrington PN, Jones B, McDowell IFW, et al. Coronary heart disease mortality in severe vs. non-severe familial hypercholesterolaemia in the Simon Broome Register. *Atherosclerosis* 2019;281:207–12. Available from: <https://doi.org/10.1016/j.atherosclerosis.2018.11.014>.
- [102] Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APHA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2018;(18):39033–8. Available from: <https://doi.org/10.1016/j.jacc.2018.11.002> pii: S0735-1097.
- [103] Soran H, Durrington P. Rosuvastatin: efficacy, safety and clinical effectiveness. *Expert Opin Pharmacother* 2008;9:2145–60.
- [104] Jones P, Kafonek S, Laurora I, Hunninghake D. Comparative dose efficacy study of atorvastatin versus simvastatin, pravastatin, lovastatin, and fluvastatin in patients with hypercholesterolemia (the CURVES study). *Am J Cardiol* 1998;81:582–7.
- [105] Banach M, Rizzo M, Toth PP, et al. Statin intolerance – an attempt at a unified definition. Position paper from an International Lipid Expert Panel. *Arch Med Sci* 2015;11:1–23.
- [106] Rosenson RS, Baker S, Banach M, et al. Optimizing cholesterol treatment in patients with muscle complaints. *J Am Coll Cardiol* 2017;70:1290–301.
- [107] Armitage J. The safety of statins in clinical practice. *Lancet* 2007;370:1781–90.
- [108] Pandor A, Ara RM, Tumur I, et al. Ezetimibe monotherapy for cholesterol lowering in 2,722 people: systematic review and meta-analysis of randomized controlled trials. *J Intern Med* 2009;265:568–80.
- [109] National Institute for Health and Clinical Excellence. Ezetimibe for treating primary heterozygous-familial and non-familial hypercholesterolaemia; 2007. Available from: <https://www.nice.org.uk/guidance/ta385>.
- [110] Serban MC, Banach M, Mikhailidis DP. Clinical implications of the IMPROVE-IT trial in the light of current and future lipid-lowering treatment options. *Expert Opin Pharmacother* 2016;17:369–80.
- [111] Cannon CP, Blazing MA, Giugliano RP, et al. Ezetimibe added to statin therapy after acute coronary syndromes. *N Engl J Med* 2015;372:2387–97.
- [112] Mazidi M, Rezaie P, Karimi E, Kengne AP. The effects of bile acid sequestrants on lipid profile and blood glucose concentrations: a systematic review and meta-analysis of randomized controlled trials. *Int J Cardiol* 2017;227:850–7. Available from: <https://doi.org/10.1016/j.ijcard.2016.10.011>.
- [113] Jaworski K, Jankowski P, Kosior DA. PCSK9 inhibitors – from discovery of a single mutation to a ground breaking therapy of lipid disorders in one decade. *Arch Med Sci* 2017;13:914–29.
- [114] National Institute of Clinical Health and Excellence (NICE). Evolocumab for treating primary hypercholesterolaemia and mixed dyslipidaemia. Technical Appraisal Guidance [TA394]. NICE: London; 2016. Available from: <https://www.nice.org.uk/guidance/ta394>.
- [115] National Institute of Clinical Health and Excellence (NICE). Alirocumab for treating primary hypercholesterolaemia and mixed dyslipidaemia. NICE: London; 2016. Available from: <https://www.nice.org.uk/guidance/ta393>.
- [116] Cicero AFG, Colletti A, Bajraktari G, et al. Lipid lowering nutraceuticals in clinical practice: position paper from an International Lipid Expert Panel. *Arch Med Sci* 2017;13:965–1005.
- [117] Buchwald H, Varco RL, Matts JP, et al. Effect of partial ileal bypass surgery on mortality and morbidity from coronary heart disease in patients with hypercholesterolemia. Report of the Program on the Surgical Control of the Hyperlipidemias (POSCH). *N Engl J Med* 1990;323:946–55.
- [118] Palacio CH, Harring TR, Nguyen NT, Goss JA, O'Mahony CA. Homozygous familial hypercholesterolemia: case series and review of the literature. *Case Rep Transplant* 2011;2011:154908.
- [119] Carswell KA, Belgaumkar AP, Amiel SA, Patel AG. A systematic review and meta-analysis of the effect of gastric bypass surgery on plasma lipid levels. *Obes Surg* 2016;26:843–55.
- [120] Yadav R, Hama S, Liu Y, et al. Effect of Roux-en-Y bariatric surgery on lipoproteins, insulin resistance, and systemic and vascular inflammation in obesity and diabetes. *Front Immunol* 2017;8:1512. Available from: <https://doi.org/10.3389/fimmu.2017.01512>.
- [121] Kennedy-Dalby A, Adam S, Ammori BJ, Syed AA. Weight loss and metabolic outcomes of bariatric surgery in men versus women – a matched comparative observational cohort study. *Eur J Intern Med* 2014;25:922–5.
- [122] Thompson GR, Catapano A, Saheb S, et al. Severe hypercholesterolaemia: therapeutic goals and eligibility criteria for LDL apheresis in Europe. *Curr Opin Lipidol* 2010;21:492–8.
- [123] Marais AD, Blom DJ. Recent advances in the treatment of homozygous familial hypercholesterolaemia. *Curr Opin Lipidol* 2013;24:288–94.
- [124] France M, Rees A, Datta D, et al. HEART UK statement on the management of homozygous familial hypercholesterolaemia in the United Kingdom. *Atherosclerosis* 2016;255:128–39.
- [125] Blom DJ, Aversa MR, Meagher EA, et al. Long-term efficacy and safety of the microsomal triglyceride transfer protein inhibitor lomitapide in patients with homozygous familial hypercholesterolemia. *Circulation* 2017;136:332–5.
- [126] McGowan MP, Tardif JC, Ceska R, et al. Randomized, placebo-controlled trial of mipomersen in patients with severe hypercholesterolemia receiving maximally tolerated lipid-lowering therapy. *PLoS One* 2012;7:e49006.
- [127] Hegele RA, Tsimikas S. Lipid-lowering agents. Targets beyond PCSK9. *Circ Res* 2019;124:386–404. Available from: <https://doi.org/10.1161/CIRCRESAHA.118.313171>.
- [128] Larsen LE, Stoekenbroek RM, Kastelein JJP, Holleboom AG. Moving targets: recent advances in lipid-lowering therapies. *Arterioscler Thromb Vasc Biol* 2019;39(3):349–59. Available from: <https://doi.org/10.1161/ATVBAHA.118.312028>.
- [129] Di Bartolo BA, Schwarz N, Andrews J, Nicholls SJ. Infusional high-density lipoprotein therapies as a novel strategy for treating atherosclerosis. *Arch Med Sci* 2017;13:210–14 [PMC free article] [PubMed] [Google Scholar].



- [130] Hovingh GK, Smits LP, Stefanutti C, et al. The effect of an apolipoprotein A-I-containing high-density lipoprotein-mimetic particle (CER-001) on carotid artery wall thickness in patients with homozygous familial hypercholesterolemia: the Modifying Orphan Disease Evaluation (MODE) study. *Am Heart J* 2015;169:736–42e1 [PubMed] [Google Scholar].
- [131] Kataoka Y, Andrews J, Duong M, et al. Regression of coronary atherosclerosis with infusions of the high-density lipoprotein mimetic CER-001 in patients with more extensive plaque burden. *Cardiovasc Diagn Ther* 2017;7:252–63.
- [132] Van Craeyveld E, Jacobs F, Gordts SC, De Geest B. Gene therapy for familial hypercholesterolemia. *Curr Pharm Des* 2011;17:2575–91.
- [133] Nikolic D, Rizzo M, Mikhailidis DP, Wong NC, Banach M. An evaluation of RVX-208 for the treatment of atherosclerosis. *Expert Opin Investig Drugs* 2015;24:1389–98.
- [134] Di Bartolo BA, Scherer DJ, Nicholls SJ. Inducing apolipoprotein A-I synthesis to reduce cardiovascular risk: from ASSERT to SUSTAIN and beyond. *Arch Med Sci* 2016;12:1302–7.
- [135] Ray KK, Landmesser U, Leiter LA, et al. Inclisiran in patients at high cardiovascular risk with elevated LDL cholesterol. *N Engl J Med* 2017;376:1430–40.
- [136] Penson P, McGowan M, Banach M. Evaluating bempedoic acid for the treatment of hyperlipidaemia. *Expert Opin Investig Drugs* 2017;26:251–9.
- [137] Group HTRC, Bowman L, Hopewell JC, Chen F, et al. Effects of anacetrapib in patients with atherosclerotic vascular disease. *N Engl J Med* 2017;377:1217–27.
- [138] Loomba RS, Arora R. Prevention of cardiovascular disease utilizing fibrates: a pooled meta-analysis. *Am J Ther* 2010;17:e182–8.
- [139] Lee M, Saver JL, Towfighi A, Chow J, Ovbiagele B. Efficacy of fibrates for cardiovascular risk reduction in persons with atherogenic dyslipidemia: a meta-analysis. *Atherosclerosis* 2011;217:492–8.
- [140] Lavigne PM, Karas RH. The current state of niacin in cardiovascular disease prevention: a systematic review and meta-regression. *J Am Coll Cardiol* 2013;61:440–6.
- [141] Group HTC, Landray MJ, Haynes R, et al. Effects of extended-release niacin with laropiprant in high-risk patients. *N Engl J Med* 2014;371:203–12.
- [142] Bhatnagar D, Durrington PN. Omega-3 fatty acids: their role in the prevention and treatment of atherosclerosis related risk factors and complications. *Int J Clin Pract* 2003;57(4):305–14.
- [143] Rizos EC, Ntzani EE, Bika E, Kostapanos MS, Elisaf MS. Association between omega-3 fatty acid supplementation and risk of major cardiovascular disease events: a systematic review and meta-analysis. *J Am Med Assoc* 2012;308:1024–33.
- [144] Sandesara PB, Virani SS, Fazio S, Shapiro MD. The forgotten lipids: triglycerides, remnant cholesterol, and atherosclerotic cardiovascular disease risk. *Endocr Rev* 2019;40(2):537–57. Available from: <https://doi.org/10.1210/er.2018-00184>.
- [145] Durrington PN, Bhatnagar D, Mackness MI, Morgan J, Julier K, Khan MA, et al. An omega-3 polyunsaturated fatty acid concentrate administered for one year decreased triglycerides in simvastatin treated patients with coronary heart disease and persisting hypertriglyceridaemia. *Heart* 2001;85(5):544–8.
- [146] Chan LN. Cardiovascular risk reduction with icosapent ethyl. *N Engl J Med* 2019;380(17):1677. Available from: <https://doi.org/10.1056/NEJMc1902165>.
- [147] Strilchuck L, Fogacci F, Cicero AF. Safety and tolerability of injectable lipid-lowering drugs: an update of clinical data. *Expert Opin Drug Saf* 2019;1–11. Available from: <https://doi.org/10.1080/14740338.2019.1620730>.

## Further reading

Gaudet et al., 2017 Gaudet D, Gipe DA, Pordy R, Ahmad Z, Cuchel M, Shah PK, et al. ANGPTL3 inhibition in homozygous familial hypercholesterolemia. *N Engl J Med* 2017;377(3):296–7. Available from: <https://doi.org/10.1056/NEJMc1705994>.





# Molecular medicine of diabetes mellitus

*Dhavendra Kumar*<sup>1,2,3</sup>

<sup>1</sup>Institute of Medical Genetics, University Hospital of Wales, Cardiff, United Kingdom <sup>2</sup>Genomic Policy Unit, University of South Wales, Pontypridd, United Kingdom <sup>3</sup>William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

## 15.1 Introduction

Diabetes mellitus (DM) is a multisystem metabolic disorder of relative insulin deficiency, resistance or both leading to persistent and irreversible chronic hyperglycemia. It is commonly discussed in the context of overweight or obesity. However, in a small proportion of cases, there is no correlation with obesity. The International Diabetes Federation estimated that 382 million people (8.3% of the global population) had diabetes in 2013 and estimates an increase to 592 million (10.1%) in 2035 [1]. There is alarming rise of prevalence in new emerging developing nations in South Asia (India), Far East (China), and Asia-Pacific (Papua New Guinea and Polynesian islands) [2]. India alone is estimated to have around 70 million overt cases and probably similar number of covert diabetics [3].

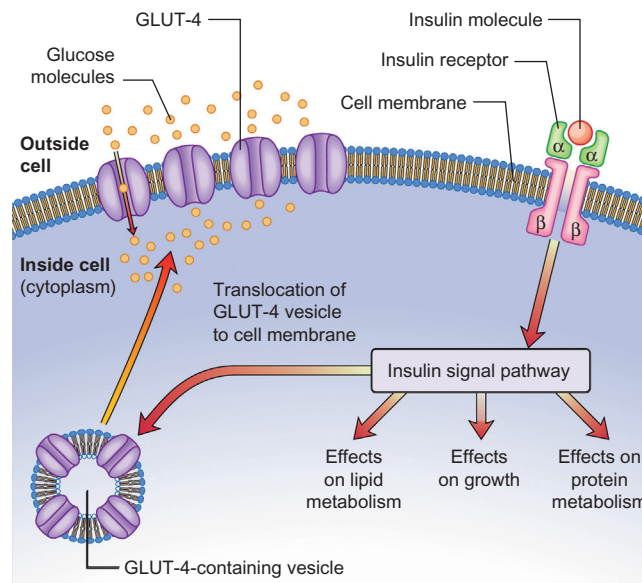
Diabetes is generally irreversible, and, although patients can lead a reasonably normal lifestyle, its late complications result in reduced life expectancy and major health costs. Associated overweight and obesity complicated long-term morbidity. Major chronic complications include micro/macrovacular disease, leading to an increased prevalence of coronary artery disease, peripheral vascular disease and stroke, and microvascular damage causing diabetic retinopathy and nephropathy. Sensory and autonomic neuropathy is another major complication [4].

This chapter describes the current understanding of the syndrome of DM in the context of molecular biology of glucose homeostasis, role of insulin secretion and resistance, genetic/genomic and molecular pathology, relationship with overweight and obesity, the spectrum of clinical manifestations, pharmacotherapy, and supportive treatment strategies for the management of DM.

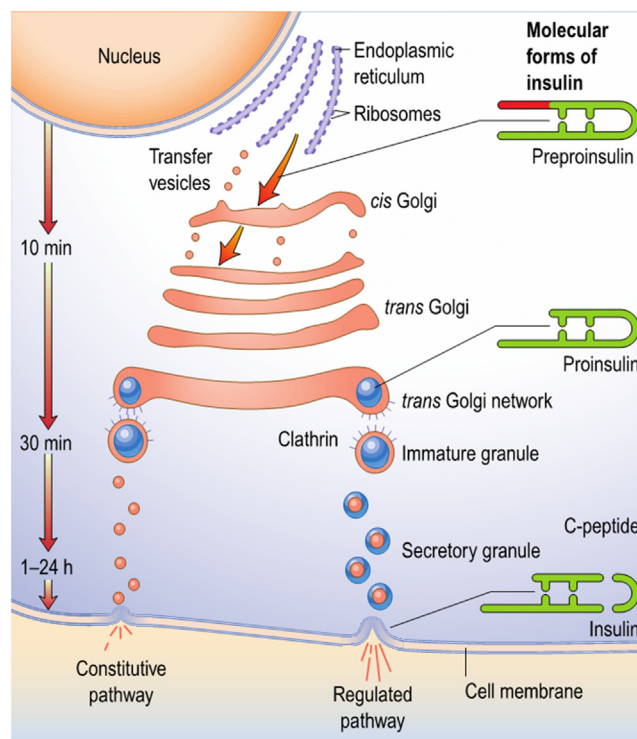
## 15.2 Molecular basis of glycemic homeostasis

Understanding molecular arrangements governing the glycemic homeostasis are essential for the diagnosis and management of DM. In normal physiological conditions the blood glucose levels are closely regulated and maintained in the range of 3.5–8.0 mmol/L (63–144 mg/dL), irrespective of food intake, fasting or physical activity. The principal organ of glucose homeostasis is the liver, which absorbs and stores glucose as glycogen in the postabsorptive state and releases it into the circulation between meals to match the rate of glucose utilization by peripheral tissues (Fig. 15.1). The liver also combines three—carbon molecules derived from the breakdown of fat (glycerol), muscle glycogen (lactate), and protein (e.g., alanine) into the six-carbon glucose molecule by the process of gluconeogenesis. On average, approximately each day 200 g of glucose is produced and consumed. The liver contributes around 90% of the glucose predominantly derived from liver glycogen and hepatic gluconeogenesis. A small proportion is also produced from renal gluconeogenesis.





**FIGURE 15.2** The role of GLUT-4 for the cellular glucose transport through stimulating the insulin signal pathway. *GLUT*, Glucose transporter. Source: Adapted from Figure 27.4 of Kumar P, Clark M. *Clinical medicine*. 9th ed., Churchill-Livingstone, Elsevier; 2017. p. 1245.



**FIGURE 15.3** The synthesis and secretion of Insulin from the  $\beta$  cells of islets of Langerhans—note constitutive and regulatory pathways. Source: Adapted from Figure 27.1 of Kumar P, Clark M. *Clinical medicine*. Churchill-Livingstone, Elsevier; 2017. p. 1244.

events triggering the release of insulin-containing granules (Fig. 15.3). The main phases of insulin secretion are preproinsulin, proinsulin, and the C-peptide-containing secretory form of insulin. After secretion, insulin enters the portal circulation and is carried to the liver, its prime target organ. About 50% of secreted insulin is extracted and degraded in the liver and kidneys remove the remaining part. C-peptide molecule of the insulin provides the measurable index of insulin secretion that is eventually degraded by the kidneys.

The insulin receptor is a glycoprotein (400 kDa), coded for on the short arm of chromosome 19. The insulin receptor is ubiquitous across the cell membrane of many cells. It consists of a dimer with two  $\alpha$ -subunits, which include the binding sites for insulin, and two  $\beta$ -subunits, which traverse the cell membrane. The insulin binds to the  $\alpha$ -subunits and induces a conformational change in the  $\beta$ -subunits, resulting in the activation of tyrosine kinase and initiation of a cascade response involving a host of other intracellular substrates. An important consequence of kinase activation is migration of the GLUT-4 to the cell surface and increased transport of glucose into the cell. The insulin-receptor complex facilitates intracellular degradation of insulin and recycled back to the cell surface.

### 15.3 Disorders of the glycemic regulation

There are several disorders etiologically related to glycemic dysregulation. A significant proportion are either known or presumed genetic and molecular diseases. A recent search on the Victor McKusick's Online Mendelian Inheritance of Man ([www.OMIM.org](http://www.OMIM.org)) using the search criteria *glycemic control or regulation* revealed 8472 and 8425 entries, respectively. Another search using *glycemic phenotypes* indicated 7770 entries. The search criteria of *glycemic family* brought up just over 16,000 entries [5].

Disorders of the glycemic regulation or control are heterogeneous with variable age at onset and pleiotropic clinical manifestations. Examples include aneuploidies (trisomy 21—Down syndrome; XXY—Klinefelter's syndrome; 45X—Turner syndrome); microdeletion syndrome (15q11—Prader–Willi syndrome and others); autosomal recessive diseases (inherited metabolic enzyme deficiency, cystic fibrosis, Friedreich's ataxia, Bardet–Biedl syndrome, etc.); autosomal dominant diseases (myotonic dystrophy, Huntington's disease, porphyria, and others); mitochondrial genetic diseases; autoimmune diseases (type 1 DM), and complex polygenic/multifactorial diseases (type 2 DM—T2D; gestational DM and obesity associated with T2D). The epigenetic mechanisms are considered to be important in the pathogenesis of few rare DM related conditions, for example, neonatal diabetes. In addition, there are several infections, neoplasia, and adverse drug responses manifesting with T2D. It is not possible to discuss these conditions in this chapter. The focus is on the syndrome of DM.

There are different distinct clinical types included within the *syndrome of DM*. In broad terms, there are two forms—primary (idiopathic) and secondary. The central factor is either quantitative insulin deficiency (T1D) or qualitative lack of action of insulin (T2D). In some cases the clinical outcome of T2D results from a combination of reduced or absent insulin action combined with lack of insulin due to progressive loss of islet of Langerhans  $\beta$  cells. A number of classification models of DM are currently used. Most clinicians and diabetologists use the American Diabetes Association (ADA) classification (Table 15.1) [6].

There are clear molecular and clinical differences between T1D and T2D (Table 15.2). It is of paramount importance that the diagnosis of these two entirely different DM disorders is clearly made for different therapeutic approaches and long-term monitoring. In practice, at the time of first diagnosis, most T1D patients are children and young adults. Rarely, slow progression of T1D may lead to late diagnosis known as *latent autoimmune diabetes in adults* (LADA). In contrast, T2D is rarely diagnosed below 40 years of age, except for uncommon syndromes complicating T2D and the maturity onset diabetes of the young (MODY). There are examples of T1D patients progressing to the clinical picture of T2D associated with overweight, obesity, and insulin receptor resistance. It is often described as “double diabetes” [7].

#### 15.3.1 Molecular mechanisms in type 1 diabetes mellitus

The type 1 DM is commonly immune-mediated, but the nonimmune-mediated form also exists [8]. The incidence of T1D is alarmingly rising, particularly in western countries at the rate of 3% annually. It belongs to a family of human leucocyte antigen (HLA)—associated immune-mediated organ-specific diseases. The mode of inheritance is polygenic/multifactorial, with the greatest contribution from the HLA region. There are many other non-HLA genes that are involved in the complex immune-mediated mechanisms. Both immune processes lead to autoantibodies directed against pancreatic islets and could appear in the circulation early in life and often predate clinical onset by many years. Older patients may have autoantibodies that appear late (LADA) or part of the multiorgan *autoimmune endocrinopathy syndrome*. All T1D patients require insulin therapy.

Increased susceptibility to T1D is inherited following the pattern of polygenic/multifactorial inheritance with around 30%–50% heritability based on concordance figures from identical twin studies. The “parent of origin”

**TABLE 15.1** The classification of diabetes mellitus.

---

*Type 1 diabetes*

- $\beta$ -Cell destruction, usually leading to absolute insulin deficiency:
  - Immune-mediated
  - Idiopathic

*Type 2 diabetes*

- May range from predominantly insulin resistance with relative insulin deficiency to a predominantly secretory defect with insulin resistance

*Other specific types*

- Genetic defects of  $\beta$ -cell function
- Genetic defects in insulin action (mainly receptor mutations)
- Diseases of the exocrine pancreas:
  - Pancreatitis
  - Trauma/Pancreatectomy
  - Neoplasia
  - Cystic fibrosis
  - Hemochromatosis
  - Fibrocalculous pancreatopathy
  - Other
- Endocrinopathies:
  - Acromegaly
  - Cushing syndrome
  - Glucagonoma
  - Pheochromocytoma
  - Hyperthyroidism
  - Somatostatinoma
  - Aldosteronoma
  - Others
- Drug- or chemical-induced:
  - Vacor (pyrinuron)
  - Pentamidine
  - Nicotinic acid (niacin)
  - $\beta$ -Blockers
  - Thyroid hormone
  - Diazoxide
  - $\beta$ -Adrenergic agonists
  - Thiazides
  - Phenytoin
  - Interferon- $\alpha$
  - Protease inhibitors
  - Immunosuppressive agents: glucocorticoids, ciclosporin, tacrolimus, and sirolimus
  - Antipsychotic agents: clozapine, olanzapine, and others
- Infections
  - Congenital rubella
  - Cytomegalovirus
  - Others

*Uncommon forms of immune-mediated diabetes*

- “Stiff person” syndrome
- Antiinsulin receptor antibodies

*Other genetic syndromes sometimes associated with diabetes*

- Down syndrome
- Friedreich’s ataxia
- Huntington’s chorea
- Laurence – Moon – Biedl syndrome
- Myotonic dystrophy
- Porphyria
- Prader – Willi syndrome
- Turner syndrome
- Wolfram syndrome
- Other

*Gestational diabetes mellitus*

---

*Adapted from American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2008; 31(Suppl 1):S55–S60.*

**TABLE 15.2** A comparison of type 1 and type 2 diabetes mellitus.

Features	Type 1	Type 2
Age	Younger (usually <30)	Older (usually >30)
Weight	Lean	Overweight
Symptom duration	Weeks	Months/years
Higher risk ethnicity	Northern European	Asian, African, Polynesian, and Native American
Seasonal onset	Yes	No
Heredity	HLA-DR3 or DR4 in >90%	No HLA links
Pathogenesis	Autoimmune disease	No immune disturbance
Ketonuria	Yes	No
Clinical	Insulin deficiency $\pm$ ketoacidosis Always need insulin	Partial insulin deficiency initially $\pm$ hyperosmolar state Needs insulin when $\beta$ cells fail over time
Biochemical	C-peptide disappears	C-peptide persists

HLA, Human leucocyte antigen.

risk of developing diabetes by the age of 20 is greater with a diabetic father (3%–7%) than with a diabetic mother (2%–3%). Earlier onset in the parent is associated with increased risk in the child. Recurrence risk of developing diabetes by the age of 20 in the second sibling of an affected child with T1DM is approximately 6%. This risk rises to about 20% in HLA-identical siblings.

### 15.3.2 Human leucocyte antigens and T1DM

The HLA genes on chromosome 6 are highly polymorphic and modulate the diverse immune system of the body. More than 90% of patients with T1D carry HLA-DR3-DQ2, HLA-DR4-DQ8, or both, as compared with some 35% of the background population. However, apart from limited predictive utility, HLA profiling does not have diagnostic advantage for T1D. All DQB1 alleles with an aspartic acid at residue 57 confer neutral to protective effects, with the strongest effect from DQB1\*0602 (DQ6), while DQB1 alleles with an alanine at the same position (i.e., DQ2 and DQ8) confer strong susceptibility. Genotypic combinations have a major influence on risk of disease. For example, HLA-DR3-DQ2/HLA-DR4-DQ8 heterozygotes have a considerably increased risk of disease. Some HLA class I alleles also modify the risk conferred by class II susceptibility genes.

### 15.3.3 Other genes or gene regions

Genome-wide association studies (GWAS) have greatly broadened our understanding of the genetic background to T1D. More than 50 non-HLA genes or gene regions associated with increased risk of T1D are currently known [8]. The HLA-associated genetic risk is also modulated by a number of other genes with small additive effect. These include the gene encoding insulin (*INS*) on chromosome 11 and a number of genes involved in immune responses, including the cytotoxic T-lymphocyte-associated protein-4 (*CTLA4*) gene, the lymphoid-specific protein tyrosine phosphatase (*PTPN22*) gene, and the interleukin (IL)-2R  $\alpha$ -subunit of the IL-2 receptor complex locus (*IL2RA*), all of which are implicated in a variety of HLA-associated autoimmune conditions [9].

### 15.3.4 Autoimmunity and type 1 diabetes mellitus

T1D is associated with other organ-specific autoimmune diseases, including autoimmune thyroid disease, coeliac disease, Addison's disease, and pernicious anemia. Histopathological evidence is available showing infiltration of the pancreatic islets by mononuclear cells. This appearance, known as insulinitis, resembles that in other autoimmune diseases such as thyroiditis. Several islet antigens have been characterized, and these include insulin itself, the enzyme glutamic acid decarboxylase, protein tyrosine phosphatase (IA-2), and the cation transporter



ZnT8. The evidence for immune-mediated pathogenesis in T1D led to the use of immunosuppressive agents such as cyclosporine that prolongs  $\beta$ -cell survival in newly diagnosed patients.

### 15.3.5 Environmental factors

The incidence of childhood T1D is rising across Europe at the annual rate of 2%–3%, suggesting that environmental factor(s) are involved in its pathogenesis. Islet autoantibodies appear in the first few years of life, indicating prenatal or early postnatal interactions with the environment. Exposures to dietary constituents, enteroviruses such as Coxsackie B4, and relative deficiency of vitamin D are some of the possible environmental triggers for T1D. A cleaner environment with less early stimulation of the immune system in childhood may reduce susceptibility for T1D similar to atopic/allergic conditions (the “hygiene hypothesis”). Conversely, more rapid weight gain in childhood and adolescence leading to increased insulin resistance might accelerate development of type 2 diabetes (the “accelerator hypothesis”). Children who test positive for two or more autoantibodies have a >80% risk of progression to T1D. The risk might approach 100% in those who additionally lose their first-phase insulin response to intravenous glucose and/or develop glucose intolerance. The ability to predict T1D with this degree of precision has opened the way to disease prevention. Further work is required to find ways of proactive preventive intervention before clinical onset of diabetes.

### 15.3.6 Type 2 diabetes mellitus

The type 2 DM (T2D) is one of the most common multifactorial complex disorders with a broad spectrum of clinical manifestations involving many organs and systems. It is the paradigm of disorders of the glycemic regulation.

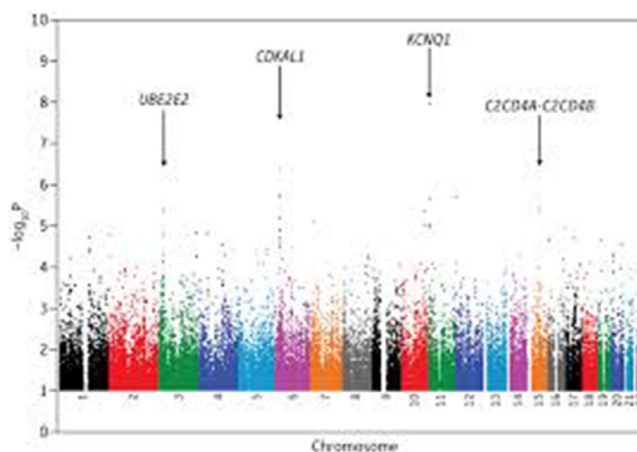
#### 15.3.6.1 Epidemiology

The epidemiology of T2D includes several factors including specific component of the genetic epidemiology. The clinical epidemiology includes several factors (Table 15.2) Since T2D is a lifelong progressive disease, many clinical and environmental factors are important from the intrauterine to late adult life [9]. Preterm birth, low birth weight as per gestational age or slow growth and development during the first year are recognized factors for the T2D pathogenesis in later life. It is unclear what is the precise mechanism(s); however, it is likely that the reduced total  $\beta$ -cell mass or relative under development of the pancreas are important contributory factors. However, studies in cystic fibrosis reveal the islet cell dysfunction is independent of the  $\beta$ -cell mass. Overweight, obesity, sedentary lifestyle, carbohydrate-rich diet, excessive alcohol consumption, mild inflammation, and possibly persistent or recurrent infection are all recognized epidemiological factors. Evidence of mild inflammation in T2D is based on mildly elevated c-reactive protein and other less common inflammatory biomarkers, particularly modestly elevated in association with raised fibrinogen and increased plasminogen activator inhibitor-1.

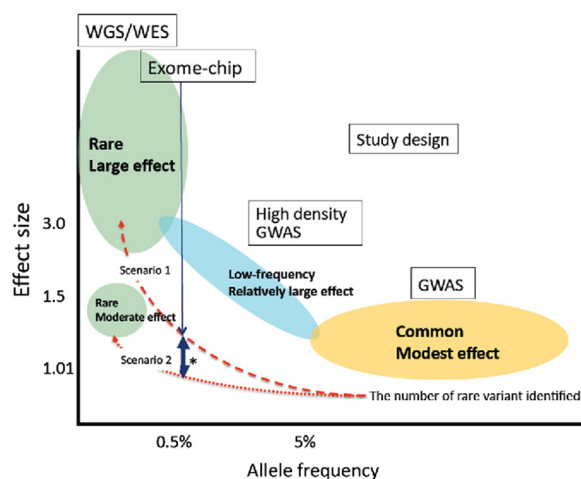
Nutritional factors are important, particularly consumption of foods with high glycemic index [10]. In keeping with the multifactorial inheritance, it is widely believed that all these environmental factors act in conjunction with genetic factors (see Section 3.2.2).

#### 15.3.6.2 Genetic factors in T2DM

Unlike T1D, several observations indicate relevance of genetic factors in T2D. Twin studies, historically, led to discussion and several studies on potential genetic factors in T2D. Identical twins of patients with T2D have more than a 50% chance of developing diabetes compared to around 25% risk to nonidentical twins or siblings. The GWAS in T2D have led to the identification of several loci that are etiologically important for the pathogenesis of T2DM [11]. Several reviews and research papers list several susceptibility loci in T2D spread across the whole human genome (Fig. 15.4) [12]. Several of these loci subserve  $\beta$ -cell development or function; however, there is no overlap with the immune function genes identified for type 1 diabetes. There is no major HLA association gene susceptibility for T2D. Other important observations on T2D gene susceptibility include, the transcription factor-7-like 2 (TCF7-L2) in Europeans, and KCNQ1 (a potassium voltage-gated channel) in Asians. TCF7-L2 carries an increased risk of around 35%, while other common variants account for no more than 10%–20%. TCF7-L2 modulates pancreatic islet cell function. So far, only few genes are known that confer increased susceptibility to T2D. In summary, apart from GWAS, several other genetic studies, including whole exome or whole genome sequencing, provide different groups of loci that are rare with large and moderate effect, common with



**FIGURE 15.4** The type 2 diabetes mellitus Manhattan chart based on GWAS [12]. GWAS, Genome-wide association studies.



**FIGURE 15.5** Outcome of large-scale genome sequencing (WES/WGS/GWAS) in common/complex medical disease [13]. GWAS, Genome-wide association studies; WES, whole exome sequencing; WGS, whole genome sequencing.

moderate effect or very mild effect (Fig. 15.5). In brief, genetic testing in T2D is limited to few rare disorders complicating with T2D or uncommon monogenic forms. In future, it is anticipated that set of few high-to-moderate risk alleles or polymorphisms could be used in triage of patients or individuals with subclinical evidence for impaired fasting glycemic control. Eventually, most public health doctors and genomic medicine clinicians would wish to apply the genome sequencing approach in selecting high-risk people within a large or moderately sized population group. However, the evidence so far rests with generous family history taking and lifestyle factors for selecting people at high risk of T2D [14].

## 15.4 Diagnosis of diabetes mellitus

### 15.4.1 Clinical manifestations

The heterogeneous syndrome of DM is a lifelong multisystem disease [15]. Clinical presentation could be acute, subacute, and chronic with recurrent acute exacerbations in poorly controlled patients. A significant proportion of cases may remain asymptomatic or nonspecific symptoms that are often ignored. Chronic clinical manifestations are likely to occur in most established diabetic individuals. Multidisciplinary diabetic centers function with the help of a number of general physicians, diabetologists, diet therapists, cardiovascular clinicians,

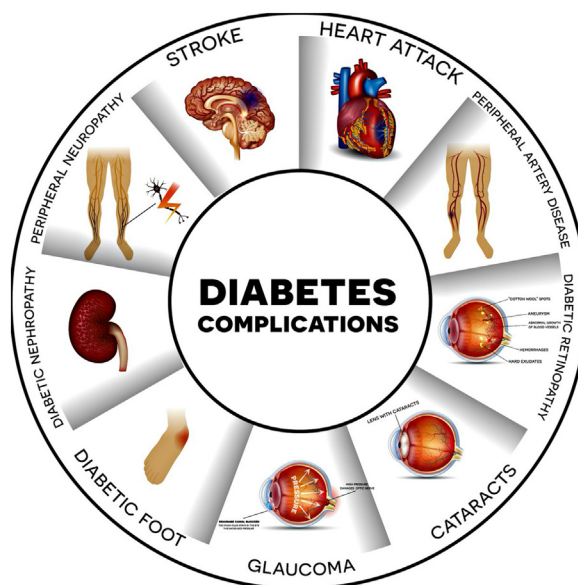


FIGURE 15.6 The multisystem complications in diabetes mellitus [15].

and few other health-care professionals. Major differences between type 1 and 2 DM are summarized in Table 15.2.

The acute presentation is typically seen in young people with a 2–6-week history and reports the classic triad of the following symptoms:

- *Polyuria* due to the osmotic diuresis resulting from increased blood glucose levels exceeding the renal threshold.
- *Thirst* due to the resulting loss of fluid and electrolytes from polyuria.
- *Weight loss* due to fluid depletion and the accelerated breakdown of fat and muscle secondary to insulin deficiency. Ketonuria is often present in young people and may progress to ketoacidosis if these early symptoms are not recognized and treated.

In subacute presentation the clinical onset may be over several months or years, particularly in older patients. The classic triad of thirst, polyuria, and weight loss may not be always present. However, patients may complain of such symptoms as lack of energy, visual blurring (owing to glucose-induced changes in refraction), or pruritus vulvae or balanitis due to *Candida* infection.

The chronic presentations are seen in most established patients. However, some of these symptoms could be the first clinical presentation that warrants full DM diagnostic profiling. These complications are multisystem (Fig. 15.6) including recurrent staphylococcal skin and fungal infections, refractive changes due to retinopathy often noted by the optician during a routine eye checkup, tingling, and numbness (paraesthesiae) due to polyneuropathy in lower limbs, erectile dysfunction, and vascular disease presenting with coronary artery disease (myocardial infarction) or signs of vascular insufficiency including gangrene.

#### 15.4.2 Blood glucose parameters—World Health Organization criteria

There are many varied diagnostic criteria for DM based on different blood glucose parameters. It is essential to use only the criteria of the World Health Organization (WHO) (Table 15.3) [16].

In most cases, fasting and random blood glucose levels are sufficient to diagnose DM. The glucose tolerance test is not routinely carried out and is useful in borderline cases and particularly in gestational DM. The term “prediabetes” is best avoided as it only indicates abnormal glucose levels within the diabetes range. Such individuals should be regarded as diabetic and managed by diet with adequate lifestyle adjustments. Diagnosis of diabetes based on glycosuria is not recommended. Any condition or old age can reduce the renal glucose threshold resulting in glycosuria. Familial renal glycosuria (OMIM 233100) is a monogenic disorder affecting function

**TABLE 15.3** The World Health Organization criteria for diabetes mellitus.

- 
- Fasting plasma glucose >7.0 mmol/L (126 mg/dL)
  - Random plasma glucose >11.1 mmol/L (200 mg/dL)
    - One abnormal laboratory value is diagnostic in symptomatic individuals
    - Two values are needed in asymptomatic people
  - Two hours after glucose tolerance test 7.8–11.0 mmol/L
    - The glucose tolerance test is only required for borderline cases and for diagnosis of gestational diabetes
- 

of the sodium–glucose cotransporter (*SGLT2*) and found in about 1:400 of the population. However, all patients presenting with glycosuria should be investigated using the WHO DM criteria [16].

### 15.4.3 HbA<sub>1c</sub> correlation

The glycosylated hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>; also referred to as A<sub>1c</sub> in the United States) is universally used as an integrated measure of an individual's past blood glucose concentrations spanning over several weeks [17]. It is accepted that an HbA<sub>1c</sub> of >6.5% (48 mmol/mol) would be considered diagnostic of diabetes, whereas a level of 5.7%–6.4% (39–46 mmol/mol) would denote increased risk of diabetes. A WHO consultation also concluded that HbA<sub>1c</sub> “can be used as a diagnostic test for diabetes.” The ADA has recommended that HbA<sub>1c</sub> should be used together with impaired glucose tolerance (IGT) and impaired fasting glycemia as a marker of “prediabetes,” with a range of 5.6%–6.4% (38–46 mmol/mol). However, the term “prediabetes” could be misleading and should be avoided.

### 15.4.4 Genetic testing

Genetic factors in T1D and T2D are discussed in the previous Section 15.3.6.2. In current practice, there is no role of genetic testing (genotyping) in the management of either form of diabetes [11]. However, specific genetic testing is recommended in certain rare syndromal forms and monogenic types of DM [18]. This decision is guided by other associated clinical features and the family history. A brief account of specific genetic tests is included in other sections of this chapter (Section 15.3.6.2).

## 15.5 Overweight, obesity, and diabetes mellitus

---

Apart from the constitutional large body size, morbid overweight and obesity deserve particular consideration due to many lifelong medical complications. It is globally endemic, particularly rapidly rising prevalence in developing and new emerging economies of the World. World statistics show an alarming rise of more than 80% during the last 20 years (Fig. 15.7) [1]. The food industry consistently gains from manufacturing and selling energy-rich food preparations and as well as many weight reducing formulas and supplements. Generous availability of fast-food chains (McDonald, KFC, Pizza Hut, etc.) is determined to waste all efforts of any personal or community weight control or reducing strategy. Arguably, the challenges associated with obtaining sustained weight reduction in those who have become overweight are a personal matter.

New emerging economies of India and China alone illustrate over 100% rise in the prevalence of morbid obesity (body mass index (BMI) >30). The risk of clinical type 2 DM (T2D) is intricately associated with obesity, largely because of extremely high risk of insulin receptor resistance and progressive loss of islet  $\beta$  cells. There are several important factors to consider for understanding the pathogenesis of T2D in relation to obesity.

### 15.5.1 Neurobiological factors

Individual neurobiological factors are particularly important that control or modify eating behavior and positive affinity for energy-dense foods. The neural conditioning of joy and physical excitement through consumption of energy-rich foods from early childhood spanning over several years leads to establishing a pattern toward being overweight and eventually the danger of morbid obesity. This pattern of eating behavior is regulated by the hypothalamus and adjoining nuclei that collectively make the biological “leptin–melanocortin” pathway governing energy intake, energy expenditure, and body weight [19]. This complex neuroendocrine hypothalamus

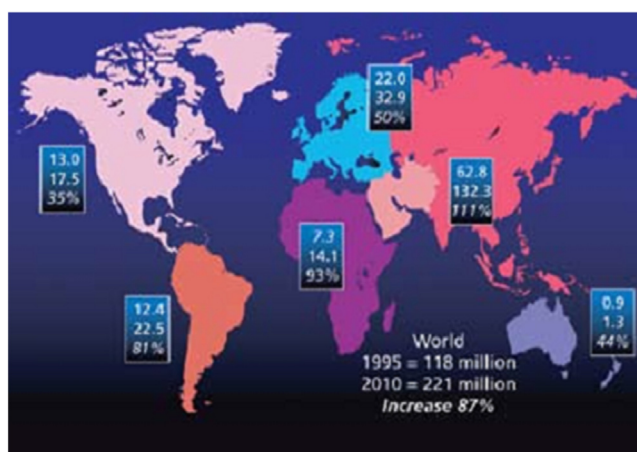


FIGURE 15.7 The global epidemic of obesity [1].

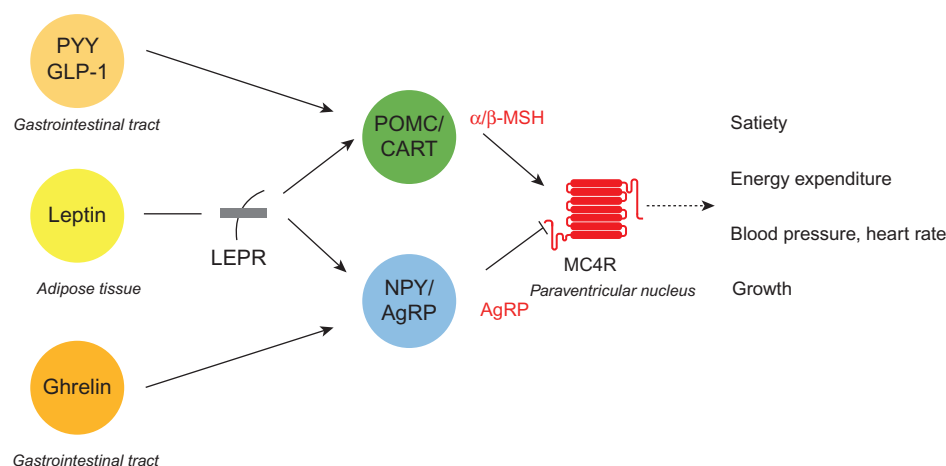


FIGURE 15.8 Melanocortin-4 receptor signalling: importance for weight regulation and obesity treatment [20]. Source: [cell.com](http://cell.com).

pathway includes the peripheral production of leptin from adipocytes. The leptin maintains a balance between the food intake and energy expenditure through series of complex steps involving a number of genes that collectively make the signaling and regulatory pathway. Major “gene-molecule” components of such pathway include NPY, AgRP, POMC, CART,  $\alpha$ -melanocyte-stimulating hormone, and melanocortin 4 receptor. In addition, orexin gene plays in regulating the energy intake and expenditure (Fig. 15.8).

In clinical practice, genetic testing for *leptin*, *leptin receptor*, *POMC*, and *MC4R* is available. Specific gene mutations help in selecting the specific drug and advising the family on recurrence risks and lifestyle modifications.

### 15.5.2 Nutritional factors—high glycemic foods

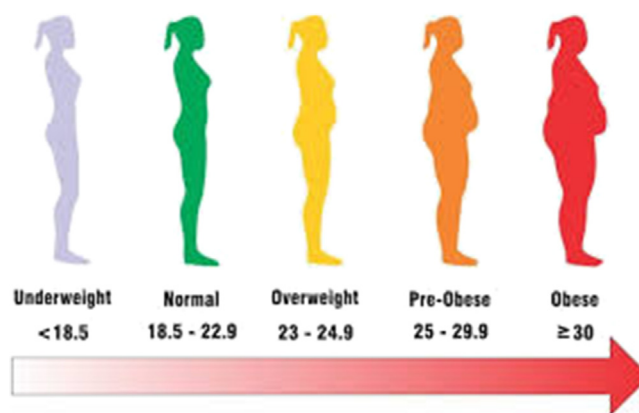
While few people are constitutionally overweight and accumulate unusually abundant adipose tissue, the main underlying factor for being overweight (BMI 23–25), mild obesity (BMI 25–29), morbid obesity (BMI 30–35), and severe obesity (BMI >35) is the personal consumption of energy-rich foods combined with relatively less or restricted physical activity.

In general, there is no disagreement on widely publicized recommendation to avoid or voluntarily restrict the carbohydrate content of food. It is debated that the best possible way would be switch over to certain food items with moderate or low glycemic index (Table 15.4). The glycemic index of any food is calculated using the guidelines by Foster-Powell and Miller [10] and expressed as percentage. In this context the clinically relevant term is the glycemic load, calculated as the glycemic index multiplied by grams of carbohydrate per serving size, indicated in parenthesis, divided by 100% [10].



**TABLE 15.4** Glycemic index and the glycemic load of selected foods [10].

Food	Glycemic index	Glycemic load
Instant rice	91	24.8 (110 g)
Baked potato	85	20.3 (110 g)
White bread	70	21.0 (two slices)
Brown (Rye) bread	65	19.5 (two slices)
Muesli	56	16.8 (110 mL)
Spaghetti	41	16.4 (55 g)
Lentil beans	29	5.7 (110 mL)
Milk	27	3.2 (225 mL)
Peanuts	14	0.7 (30 g)
Apple	36	8.1 (170 g)
Banana	53	13.3 (170 g)

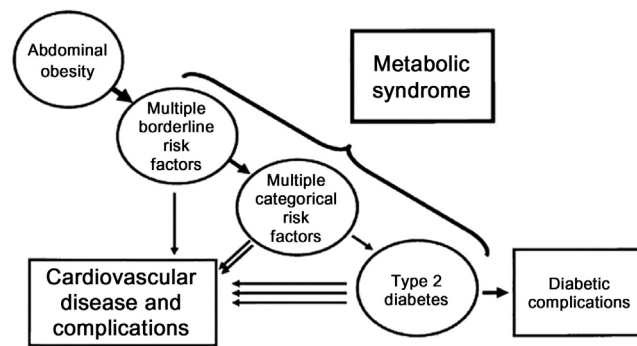
**FIGURE 15.9** Distribution of weight patterns and obesity.

### 15.5.3 Constitutional and medical obesity

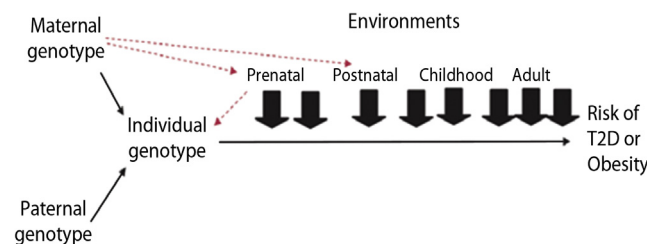
The societal perception of overweight and obesity is generally negative, particularly for and among the females. This is unacceptable, as many such people happen to be overweight and phenotypically obese simply due to the constitution and body habitus. In most cases, there is no underlying genetic or medical cause. The family history is likely as many overweight and obese people share similar dietary habits and lifestyle, often common to many people and families within a community. It is important that simply being overweight or obese does not become a stigma manifesting with unnecessary psychosocial consequences. However, it is important for these people to be on guard due to low threshold for medical complications, notably T2D and cardiovascular problems [21] (Fig. 15.9).

In the context of obesity related medical complications, the metabolic syndrome is widely discussed, however poorly understood. Typically, the metabolic syndrome refers to an individual of both genders, around 40–50 years of age, with central distribution of fat and evidence of either subclinical (impaired fasting hyperglycemia) or clinically overt T2D. Associated cardiovascular complications, particularly systemic hypertension, are common (Fig. 15.10). The condition, with its many forms, is undoubtedly heterogeneous and late outcome of several consequential factors including parental genotypes and environmental interactions beginning in early fetal life to adulthood (Fig. 15.11). The risk of T2D is alarmingly very high. From the public health perspectives, this disorder provides an example where multitude of lifelong interventions might help reducing the incidence and prevalence of the metabolic syndrome and thus prevention of T2D [22].





**FIGURE 15.10** Obesity and diabetes mellitus (T2D) closely linked to the metabolic syndrome [22].



**FIGURE 15.11** Gene–environment interaction of parental genotypes and lifetime exposures in the pathogenesis of obesity and type 2 diabetes mellitus [23].

### 15.5.4 Genetic/genomic factors

Numerous studies are on record for elucidating the genetic basis of obesity. There are many genetic and genomic forms of obesity, such as syndromic monogenic, nonsyndromic monogenic, oligogenic, and polygenic obesity [24]. Rare genetic conditions associated with overweight and obesity are discussed in a previous chapter on obesity and overweight. Recent studies, applying the genome-based laboratory methods, have accumulated evidence for the role of epigenetics/epigenomics and metagenomics in the complex pathogenesis of obesity.

#### 15.5.4.1 Family history and heritability

The family history of obesity or constitutional overweight in close or distant relatives is common. From simple observations or systemic studies the heritability is consistently between 40 and 50. It is similar to T2D. Apart from sharing the Mendelian proportion of the parental genotypes, associated factors, such as restricted fetal growth, low birth weight, childhood nutritional problems, dietary factors (rich in fat and carbohydrates), lack of physical activity, sedentary occupation, etc.), are contributory to the eventual obesity phenotype, quantitatively labeled mild, moderate, morbid, or very severe commonly using the BMI scale. Other parameters, such as abdominal girth compared to the hip (hip–waist ratio), are important but are less commonly used in clinical practice. It is mandatory for any such person to be monitored for T2D with a low threshold for medical prophylaxis with biguanide oral antidiabetic drugs (e.g., Metformin).

#### 15.5.4.2 Rare monogenic diseases and syndromes of obesity

There are several uncommon monogenic, oligogenic, and chromosomal (genome architecture) diseases recognized with the obesity phenotype. Some of these are discussed in a previous chapter (Chapter ...). Some of these illustrate key cellular and molecular mechanisms, for example, ciliopathies (Bardet–Biedl syndrome) and laminopathies (Obesity, T2D, and polycystic ovary syndrome). At present, 10 genetically elucidated obesity syndromes are known, some of which are involved in ciliary functioning. The OMIM lists 11 monogenic obesity genes involved in energy maintenance as part of the leptin–melanocortin pathway.

### 15.5.4.3 Polygenic/multifactorial obesity—genome-wide association studies /copy number variants/single-nucleotide polymorphisms

With the emergence of GWAS over the last decade, 227 genetic variants involved in different biological pathways (central nervous system, food sensing and digestion, adipocyte differentiation, insulin signaling, lipid metabolism, muscle and liver biology, and gut microbiota) have been associated with polygenic obesity.

The evidence from many GWAS in obesity is consistently in favor of polygenic/multifactorial causation. The obesity genome map illustrates around 250 loci with low-to-medium risk predisposition for obesity (Fig. 15.12) [25].

The obesity predisposition genome variants include single-nucleotide polymorphisms, copy number variants, and specific genes (POMC/MC4R). There are several ways by which a variant could manifest with the obesity phenotype. Clearly, variants collectively impact on or interfere with a number of critical gene-molecular pathways. In contrast, mutations in a specific gene, likely to be part of gene-molecular family, manifest specifically interfering with series of cellular functions leading to a defined molecular pathology manifesting with obesity (Fig. 15.13).

### 15.5.4.4 Environment and epigenetics/epigenomics

In previous sections, nutritional and lifestyle factors are elucidated contributing to the obesity phenotype. In addition, there are other factors that operate outside the functional parts of the genome. This phenomenon is discussed under epigenetics and epigenomics. There are several mechanisms as discussed elsewhere (Chapter ...). The parental genetic imprinting is one of the epigenetic mechanisms, for example, lack paternal contribution

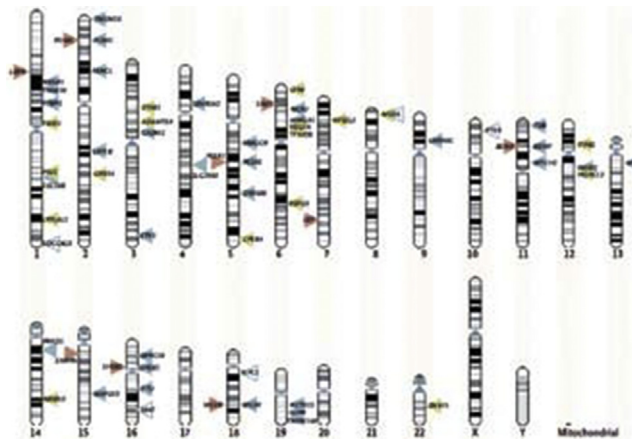


FIGURE 15.12 The obesity genome map showing critical genes and loci across the whole human genome [25].

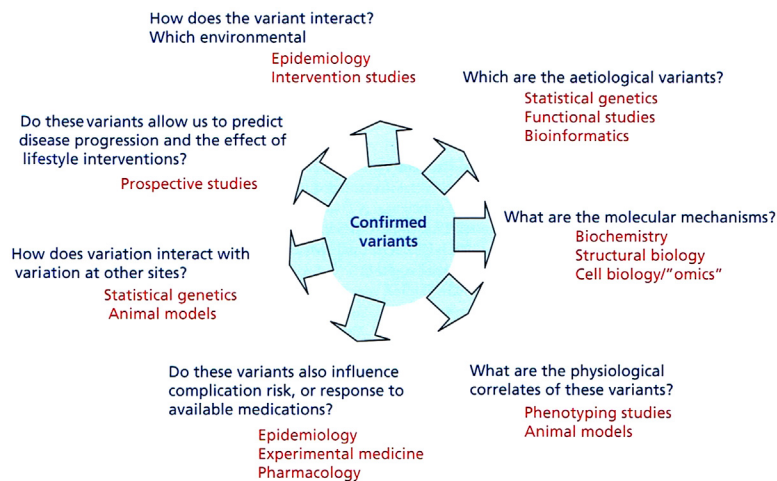


FIGURE 15.13 The clinical outcomes of obesity and diabetes genomic research [26].

(maternal disomy for 15q11) is critical for obesity in Prader–Willi syndrome. In addition, there are likely to be many other epigenetically important loci probably important for obesity. Mapping of the methylation genome (methylome) might offer better clarification of the role of epigenomics in the causation of obesity.

Advances in obligatory and facilitated epigenetic variation, and gene–environment interaction studies have partly accounted for the missing heritability of obesity and provided additional insight into its etiology [27].

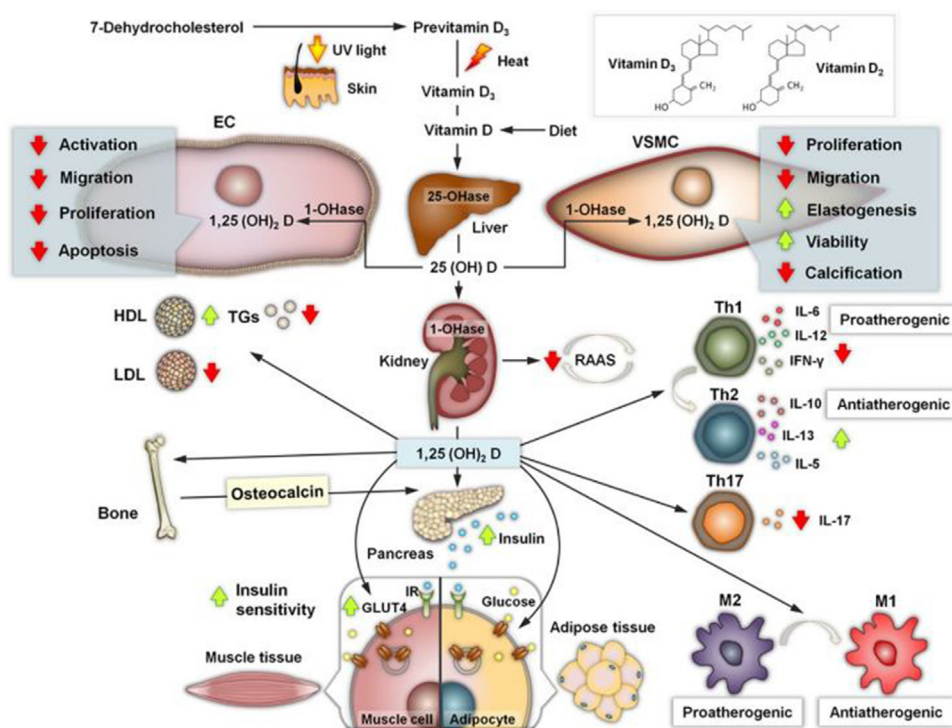
#### 15.5.4.5 Metagenomics and obesity

Being a complex and heritable disorder, obesity results from the interplay between genetic susceptibility, epigenetics, metagenomics, and the environment [27]. Large-scale alterations of the gut microbiota and its microbiome (gene content) are associated with obesity and are responsive to weight loss. Gut microbes can impact host metabolism via signaling pathways in the gut, with effects on inflammation, insulin resistance, and deposition of energy in fat stores [28].

In brief, evidence so far points to the pattern of gut bacterial flora associated with additional energy input contributory to obesity. In contrast, lean persons harbor many gut microbial that restrict energy extraction and assimilation. However, clinical correlation of gut microbial flora patterns (metagenomics) with either weight gain or loss is unclear. Further studies are warranted on the role of metagenomics in relation to obesity.

### 15.6 Vitamin D and diabetes mellitus

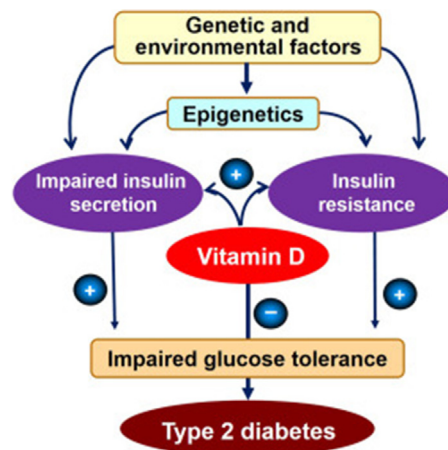
The conventional teaching and clinical practice for vitamin D deficiency are limited to bone and calcium homeostasis with the clinical hallmark of rickets. However, the vitamin D deficiency, whether nutritional or lack of sunlight exposure, has implications in many other physiological changes. Molecular studies in uncommon inherited forms of vitamin D dependent or resistant rickets facilitated understanding the complex molecular biology of vitamin D. The biologically active form of vitamin D, the  $1,25(\text{OH})_2\text{D}_3$ , results from the final enzymatic action of  $1-\alpha$  hydroxylase enzyme in the kidney (Fig. 15.14). This active form of vitamin D acts through vast network of receptors.



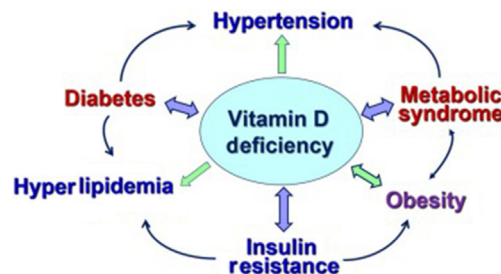
**FIGURE 15.14** Molecular regulation of metabolic functions of vitamin D. Source: Adapted with permission Wimalawansa SJ. Role of vitamin D with insulin resistance, obesity, type 2 diabetes, and metabolic syndrome. *J Steroid Biochem Mol Biol* 2018;175:177–89. <https://doi.org/10.1016/j.jsbmb.2016.09.017>.

Multitude of functions of vitamin D is mediated by ubiquitously distributed intracellular vitamin D receptors facilitated by the 1- $\alpha$  hydroxylase enzyme. These are distributed in all tissues suggesting a multitude of functions of vitamin D. There is now ample evidence that vitamin D plays an indirect but an important role in carbohydrate and lipid metabolism as reflected by its association with type 2 DM (T2D), metabolic syndrome, insulin secretion, insulin resistance, polycystic ovarian syndrome, and obesity. In addition to the widely accepted physiological musculoskeletal effects, vitamin D decreases the insulin resistance, severity of T2D, asymptomatic diabetes, metabolic syndrome, inflammation, and autoimmunity. Vitamin D exerts autocrine and paracrine effects such as direct intracellular effects via its receptors and the local production of 1,25(OH) $_2$ D $_3$ , especially in muscle and pancreatic  $\beta$ -cells. It also regulates calcium homeostasis and calcium flux through cell membranes, and activation of a cascade of key enzymes and cofactors associated with metabolic pathways (Fig. 15.14) [29].

Several studies provide evidence of inverse relationship of vitamin D status with hyperglycemia and glycemic control in patients with T2D, decrease the rate of conversion of covert diabetes to overt diabetes, and obesity. It is postulated that vitamin D acts with other genetic factors including epigenetic factors on the insulin receptors and thus the IGT (Fig. 15.15). On this basis, vitamin D appears to have indirect role in the development of obesity and hypertension, collectively recognized as the metabolic syndrome (Fig. 15.16). While a large number of observational studies support improving T2D, insulin resistance, obesity, and metabolic syndrome with vitamin D supplementation, there is, however, a lack of conclusive evidence from randomized control clinical trials that these disorders are prevented the following optimization of serum levels of 25(OH)D. Nevertheless, the scientific and molecular basis of the role of vitamin D in glycemic control, insulin resistance, and probably metabolic syndrome justify gathering further evidence through carefully designed new clinical trials. Many physicians, including diabetologists, remain in favor of including vitamin D $_3$  preparations (cholecalciferol) in managing chronic T2D.



**FIGURE 15.15** The role of vitamin D in the genetic pathology of type 2 diabetes. Source: Adapted with permission Wimalawansa SJ. Role of vitamin D with insulin resistance, obesity, type 2 diabetes, and metabolic syndrome. *J Steroid Biochem Mol Biol* 2018;175:177–89. <https://doi.org/10.1016/j.jsmb.2016.09.017>.



**FIGURE 15.16** Correlation of vitamin D deficiency with the insulin resistance and pathogenesis of T2D including hypertension and obesity. Source: Adapted with permission Wimalawansa SJ. Role of vitamin D with insulin resistance, obesity, type 2 diabetes, and metabolic syndrome. *J Steroid Biochem Mol Biol* 2018;175:177–89. <https://doi.org/10.1016/j.jsmb.2016.09.017>.



## 15.7 Inherited monogenic diabetes mellitus

In this section, rare monogenic forms of the  $\beta$ -cell dysfunction are discussed. The main clinical subtypes include neonatal diabetes, MODY, and as a manifestation of multisystem genetic syndrome [18]. DM is also a recognized clinical manifestation of the multisystem mitochondrial genetic disorders.

### 15.7.1 Neonatal diabetes mellitus

Neonatal presentation of DM (NDM) is uncommon with a prevalence of around 1 in 200,000. It is essentially type 1 diabetes (T1D) and presents early during first 6 months of life compared to the classic autoimmune T1D that never presents before 6 months of age. Apart from few familial autosomal recessive cases, most NDM cases are de novo. It is a heterogeneous condition and presents either as transient neonatal DM (TNDM) or permanent neonatal DM (PNDM) [30]. Genetic testing is available and can confirm the diagnosis in around 80% cases ([www.diabetesgenes.org](http://www.diabetesgenes.org)). Rare form of X-linked recessive neonatal diabetes is also described with immunodysregulation, polyendocrinopathy and enteropathy, the IPEX syndrome (OMIM 304790).

#### 15.7.1.1 Transient neonatal diabetes mellitus (OMIM 601410)

It is a very rare but a distinct form of NDM. Typically the neonate presents with growth delay with markedly elevated blood glucose in the first week of life. Most cases progress to remission within 3–12 months of age. However, few children may relapse later in life with noninsulin type DM. It is an imprinted disorder with over-expressed genes at the 6q24 locus. Molecular epigenetic mechanisms include paternal uniparental disomy of chromosome 6, paternally inherited microduplication of 6q24, or a methylation mutation due to inactivation of *ZFP57* or activating mutations in *KCNJ11* or *ABCC8* genes. Treatment is by short acting insulin with good clinical outcomes [31].

#### 15.7.1.2 Permanent neonatal diabetes mellitus (OMIM 606176)

Most patients are diagnosed within the first 6 months of life, and diabetes persists for longer variable period. In around 40%–50% of the cases, mutations in the  $\beta$ -cell KATP channel genes (*KCNJ11* and *ABCC8*) are etiologically linked with PNDM [32]. There are around 16 different known genes associated with PNDM. Genetic confirmation of the specific type is important for effective management by sulfonylurea tablets. The majority of patients perform very well with chlorpropamide or any other oral sulfonylureas (glibenclamide, gliclazide, etc.) preparations.

### 15.7.2 Maturity onset diabetes of the young (OMIM 125850)

It is a distinct but uncommon form of the autosomal dominant monogenic inherited DM. The majority of patients (under 25 years) typically present with variable family history in keeping with the autosomal dominant inheritance. Later presentation of this condition is also known indicating low threshold for making this diagnosis. It behaves like T2D but occasionally the disorder is mistaken as T1D. There are several types of MODY (Table 15.5—classification of MODY) with distinct gene mutations [33]. In most cases of MODY, three genes are encountered such as *glucokinase*, *HNF1A*, and *HNF4A*; mutations in *HNF1B* associated T2D also manifest with renal cysts and other genital tract developmental difficulties (RCAD syndrome). MODY is essentially a T2D disorder, however, not insulin dependent. Some cases are very slim and present with marked hyperglycemia, often mistaken as T1D. These patients might require insulin for treatment. Mutations in the *glucokinase* gene (*GCK*) cause mild, stable fasting hyperglycemia (5.5–8 mmol/L); these patients do not require any form of pharmacological treatment [34]. Individuals with the *HNF1A* mutations are very sensitive to sulfonylureas and would require careful dosage and side effects monitoring.

Genes for monogenic forms of DM include the MODY type [18]. The *glucokinase* gene is intimately involved in the glucose-sensing mechanism within the pancreatic  $\beta$  cell. The hepatic nuclear factor (*HNF*) genes and the insulin promoter factor-1 gene control nuclear transcription in the  $\beta$  cell, where they regulate its development and function. Abnormal nuclear transcription genes may cause pancreatic agenesis or progressive pancreatic damage. A handful of families with autosomal dominant diabetes have been described with mutations in neurogenic differentiation factor-1.

**TABLE 15.5** Genes for monogenic (Mendelian) disorders of the  $\beta$  cell.

Features	<i>HNF4A</i>	Glucokinase	<i>HNF1a</i>	<i>IPF-1</i>	<i>HNF1b</i>
Chromosomal location	20q	7p	12q	13q	17q
Proportion of all cases	5%	15%	70%	<1%	2%
Onset	Teens/Thirties	Present from birth	Teens/Twenties	Teens/Thirties	Teens/Twenties
Progression	Progressive hyperglycemia	Little deterioration with age	Progressive hyperglycemia	Progression unclear	Progression unclear
Microvascular complications	Frequent	Rare	Frequent	Few data	Frequent
Other features	None	Reduced birth weight	Sensitivity to sulfonylureas	Pancreatic agenesis in homozygotes	Renal cysts, proteinuria, chronic kidney disease

### 15.7.3 Mitochondrial diabetes mellitus

The pancreatic  $\beta$ -cell dysfunction is a recognized manifestation of abnormal mitochondrial function (see Chapter 1, The human genome and molecular medicine). The most common condition complicating diabetes is mitochondrial sensorineural deafness caused by mtDNA 3243A > G mutation. This mutation is also found in other condition, referred to as MELAS (myoclonic epilepsy with lactic acidosis and stroke like symptoms). This is exclusively maternally inherited. It is important to evaluate the family history for more similarly affected patients [35].

### 15.7.4 Syndromes of inherited insulin resistance

There are distinct rare monogenic disorders with inherited insulin resistance [5]. Apart from early onset type 2 DM, there are other distinguishing clinical manifestations. Specific genetic testing is recommended supported by genetic counseling.

Homozygous or compound heterozygous mutations cause severe insulin resistance and extremely high levels of insulin levels. Inheritance is usually autosomal recessive. Parental consanguinity is likely and should be clarified.

Leprechaunism or Donohue syndrome is the most severe form of insulin resistance and presents in early infancy manifesting with pre/postnatal failure to thrive, hirsutism, progeria such as facies, thick lips with prominent ears, and enlargement of the breasts and genitalia.

Rabson–Mendenhall syndrome is characterized with acanthosis nigricans, polycystic ovaries, and virilization of females. It is less severe form and usually presents later. Inheritance is autosomal recessive.

Lipodystrophy with insulin resistance are uncommon genetic conditions. These syndromes are characterized with complete or partial loss of subcutaneous adipose tissue, hyperglycemia, hyperlipidemia, and diabetes. The familial partial lipodystrophy is an autosomal dominant disorder caused by missense mutations in lamin A/C gene (*LMNA*). The loss of adipose tissue is most marked in extremities and trunk with relatively increased fatty tissue accumulation in head and neck regions. In contrast to *LMNA* mutations, other genetic defects cause autosomal recessive (AR) type of congenital generalized lipodystrophy caused by mutations in the *BSCL2* or *AGPAT2* genes.

### 15.7.5 Malformation syndromes with diabetes mellitus

There are few rare genetic syndromes that are known to complicate DM. These disorders are discussed in earlier chapter and summarized in this section (Table 15.6).



**TABLE 15.6** Associations of diabetes with genetic syndromes.

Deafness	Eye signs	Renal disease	Severe obesity	Insulin resistance
MELAS	Prader–Willi	RCAD	Prader–Willi	Donohue
Wolfram	Alstrom	Wolfram	Alstrom	Rabson–Mendenhall
	Bardet–Biedl	Bardet–Biedl	Bardet–Biedl	Partial lipodystrophy
	Wolfram			

## 15.8 Management of diabetes mellitus

Management of DM requires patient-led multidisciplinary team care. There are several information resources available on the management of diabetes that is broadly similar. Details on this aspect of diabetes are beyond the scope of this chapter. Major components of diabetes care and management are summarized. The management of acute (hypoglycemia, ketoacidosis, and acute renal failure) and lifelong chronic complications of diabetes (hypertension, retinopathy, nephropathy, polyneuropathy, and peripheral vascular disease) are not discussed in this chapter.

### 15.8.1 Diet/nutritional supplements

Dietary management is the key and often ignored or given less importance. Interested reader or student is advised to check dietary protocol used in a nearby diabetic clinic. In general, typical diet for DM patient should be

- low in sugar but does not need to be sugar-free,
- high in starchy carbohydrate (especially foods with a low glycemic index), that is, slower absorption,
- high in fiber, and
- low in fat (especially saturated fat).

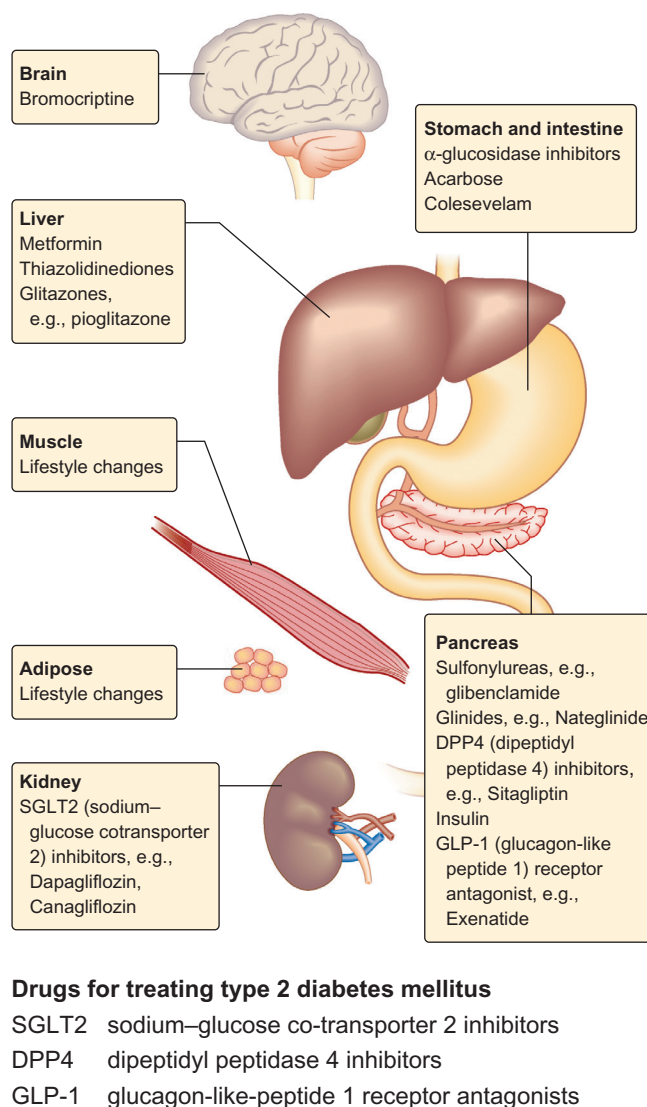
Other nutritional supplements, particularly vitamin D, are also recommended. However, the evidence is not overwhelming due to limited information available. In general, the salt intake should also be kept low due to high risk of coexisting hypertension and cardiovascular risk. In general, dietary management should be combined with generous mild to moderate physical activity.

### 15.8.2 Insulin

Insulin is the mainstay for acute and long-term treatment of T1D and in some instances of T2D. Any medical emergency in a diabetic patient should be managed by insulin injection or infusion. Most poorly controlled T2D patients require combined insulin and oral antidiabetic drugs. There are many forms of insulin preparations are now available, broadly classified as short, intermediate, and long acting. All forms of insulin are now recombinant biosynthetic manufactured by using bacterial or fungal clones. There are now several molecular modifications of insulin analogues that are regularly prescribed to both T1D and T2D patients. Details on specific insulin choice and clinical pharmacology are not discussed here. It is important to discuss this aspect of care with a specialist physician or dedicated diabetic clinic.

### 15.8.3 Oral antidiabetic drugs

Large majority of patients with T2D are managed with one or more oral antidiabetic drugs. Apart from drugs acting on insulin receptors (biguanides) or stimulating the  $\beta$  cells (sulfonylureas), new classes of drugs are now available (Fig. 15.17). The glycosylated hemoglobin (Hb1AC) levels are used for individual patient specific selection of the oral antidiabetic drug (Fig. 15.18).



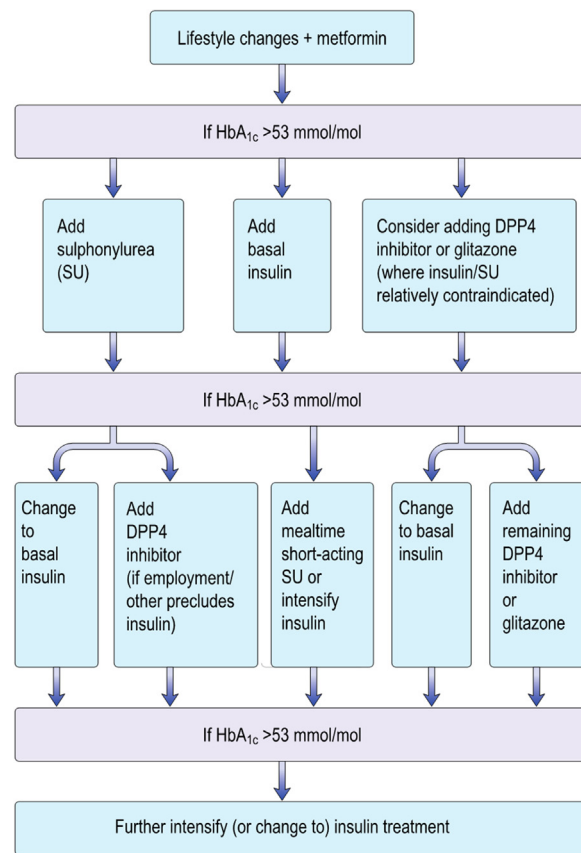
**FIGURE 15.17** Commonly used oral antidiabetic drugs; note organ specific site of action.

Selection of oral antidiabetic drugs based on HbA1C levels. Criteria for including new drugs, such as GLP-1 agonists, sodium glucose cotransporter 2 inhibitors, and dipeptidyl peptidase-4, are shown. All T2D patients would also require daily low-dose aspirin, statin therapy, and adequate antihypertensive control.

## 15.9 Summary

DM is a global phenomenon with huge health and socioeconomic implications. It is endemic in both developed and most developing countries with new emerging economies. It mirrors the societal changes and lifestyle as is evident from alarmingly high prevalence across the Indian subcontinent with millions more added every year. Apart from the health burden of infectious communicable diseases, the DM is by far the most important disease in the noncommunicable diseases group. To large extent the disease is preventable, and with early diagnosis and medical interventions, most severe and unpleasant complications could be avoided.

Main clinical syndromes of DM include type 1 and type 2 forms with clearly defined genetic and molecular pathologies. Some of these are now applied in early diagnosis and guiding the antidiabetic therapy. In addition, there are other forms of DM, specifically the MODY, TNDM, and PNDM, rare inherited disorders of the insulin resistance, and uncommon malformation syndromes complicated with DM. Specific genetic testing for these



**FIGURE 15.18** Recommended clinical pathway for type 2 diabetes mellitus. Source: Adapted with permission from Figure 21.10 of Kumar P, Clark M. Chapter 27—Diabetes mellitus. In: *Clinical medicine*, 9th ed. Churchill-Livingstone, Elsevier; 2017. p. 1256.

syndromes is available. Role of clinical genetics and genetic counseling is acknowledged and needs to be included within the multidisciplinary management of DM.

Apart from various forms of DNA recombinant forms of insulin, several oral antidiabetic drugs are now available, conveniently groups as first, second, and third line with few next-generation drugs on the horizon. Monitoring of oral DM therapy with any drug requires periodic hyperglycemia surveillance using the glycosylated hemoglobin (HbA<sub>1c</sub>). Surgical intervention might be necessary, specifically to manage progressive vascular and neurological complications.

## References

- [1] Zimmet P, Alberti KGMM, Shaw J. Global and societal implications of the diabetes epidemic. *Nature* 2001;414:782–7.
- [2] Yang W, Lu J, Weng J. Prevalence of diabetes among men and women in China. *N Engl J Med* 2010;362:1090–101.
- [3] Diabetes. Diabetes in India. Diabetes, UK. <<https://www.diabetes.co.uk/global-diabetes/diabetes-in-india.html>>.
- [4] Tuoml T, Santoro N, Caprio S, et al. The many faces of diabetes, a disease with increasing heterogeneity. *Lancet* 2014;383:1081–97.
- [5] Victor McKusick's online Mendelian inheritance in man. <[www.OMIM.org](http://www.OMIM.org)>.
- [6] American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2008;31(Suppl 1):S55–60.
- [7] Randle PJ, Garland GP, Hales CN, Newshome EA. The glucose fatty-acid cycle. Its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus. *Lancet* 1965;1:785–9.
- [8] Bennett ST, Todd JA. Human type 1 diabetes and the insulin gene: principles of mapping polygenes. *Annu Rev Genet* 1996;30:343–70.
- [9] Smyth D, Cooper JD, Collins JE, Heward JM, Franklyn JA, Howson JM, et al. Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes, and evidence for its role as a general autoimmunity locus. *Diabetes* 2004;53:3020–3.
- [10] Foster-Powell K, Miller JB. International tables of glycemic index. *Am J Clin Nutr* 1995;62(4):871S–90S.
- [11] McCarthy MI. Genomics, type 2 diabetes, and obesity. *New Engl J Med* 2012;363(24):2339–50.
- [12] Voight BF, Scott LJ, Steinthorsdottir V, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010;42:579–89.
- [13] Manolio TA, Collins FS, Visscher PM, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.

- [14] Valdez R, Yoon PW, Qureshi N, Green RF, Khoury MJ. Family history in public health practice: a genomic tool for disease prevention and health promotion. *Annu Rev Public Health* 2010;31(1):69–87.
- [15] Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* 2005;365:1333–46.
- [16] Alberti KGM, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. 1. Diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998;15:539–53.
- [17] American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014;37(Suppl 1):S81–90. Available from: [http://www.who.int/diabetes/publications/WHO\\_report:\\_use\\_of\\_glycated\\_haemoglobin\\_\(HbA1c\)\\_in\\_the\\_diagnosis\\_of\\_diabetes\\_mellitus](http://www.who.int/diabetes/publications/WHO_report:_use_of_glycated_haemoglobin_(HbA1c)_in_the_diagnosis_of_diabetes_mellitus).
- [18] McCarthy MI, Hattersley AT. Molecular diagnostics in monogenic and multifactorial forms of type 2 diabetes. *Expert Rev Mol Diagn* 2001;1:403–12.
- [19] Larsen LH, Echwald SM, Sørensen TIA, Andersen T, Wulff BS, Pedersen O. Prevalence of mutations and functional analyses of melanocortin 4 receptor variants identified among 750 men with juvenile-onset obesity. *J Clin Endocrinol Metab* 2005;90:219–24.
- [20] Kühnen P, Krude H, Biebermann H. Melanocortin-4 Receptor Signalling: Importance for Weight Regulation and Obesity Treatment. *Trends in Molecular Medicine* 2019;25(2):136–48.
- [21] Kopelman P. Obesity as a medical problem. *Nature* 2000;404:635–43.
- [22] Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet* 2005;365:1415–28.
- [23] Walley AJ, Blakemore AIF, Froguel P. Genetics of obesity and the prediction of risk for health. *Hum Mol Genet* 2006;15(Suppl 2):R124–30.
- [24] Bell CG, Walley AJ, Froguel P. The genetics of human obesity. *Nat Rev Genet* 2005;6:221–34.
- [25] Meyre D, Delplanque J, Chèvre JC, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat Genet* 2009;41:157–9.
- [26] Shuldiner AR, Munir KM. Genetics of obesity: more complicated than initially thought. *Lipids* 2003;38(2):97–101.
- [27] Pigeayre M, Yazdi FT, Kaur Y, Meyre D. Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clinical Science* 2016;130(12):943–86. Available from: <https://doi.org/10.1042/CS20160136>.
- [28] Ley RE. Obesity and the human microbiome. *Curr Opin Gastroenterol* 2010;26(1):5–11. Available from: <https://doi.org/10.1097/MOG.0b013e328333d751>.
- [29] Wimalawansa SJ. Role of vitamin D with insulin resistance, obesity, type 2 diabetes, and metabolic syndrome. *J Steroid Biochem Mol Biol* 2018;175:177–89. Available from: <https://doi.org/10.1016/j.jsbmb.2016.09.017>.
- [30] Aguilar-Bryan L, Bryan J. Neonatal diabetes mellitus. *Endocr Rev* 2008;29(3):265–91. Available from: <https://doi.org/10.1210/er.2007-0029> Epub 2008 Apr 24.
- [31] Ma D, Shield JP, Dean W, Leclerc I, Knauf C, Burcelin RR, et al. Impaired glucose homeostasis in transgenic mice expressing the human transient neonatal diabetes mellitus locus, TNDM. *J Clin Invest* 2004;114:339–48.
- [32] Gloyn AL, Pearson ER, Antcliff JF, Proks P, Bruining GJ, Slingerland AS, et al. Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. *New Eng J Med* 2004;350:1838–49 Note: Erratum: *New Eng. J. Med.* 351: 1470 only, 2004.
- [33] Fajans SS, Bell GI, Polonsky KS. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N Engl J Med* 2001;345:971–80. Available from: <https://doi.org/10.1056/NEJMra002168>.
- [34] Njolstad PR, Søvik O, Antonio C-M, Bjørkhaug L, Massa O, Barbetti F, et al. Neonatal diabetes mellitus due to complete glucokinase deficiency. *N Engl J Med* 2001;344(21). Available from: [www.nejm.org](http://www.nejm.org).
- [35] Kadowaki T, Kadowaki H, Mori Y, Tobe K, Sakuta R, Suzuki Y, et al. A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. *New Eng J Med* 1994;330:962–8.

## Further reading

- Bener A, Zirie M, Al-Rikabi A. Genetics, obesity, and environmental risk factors associated with type 2 diabetes. *Croat Med J* 2005;46(2):302–7.

# Molecular genetic management of epilepsy

David Lewis-Smith<sup>1,2</sup>, Donald P. Craig<sup>1</sup>, Naomi J.P. Thomas<sup>3</sup>,  
Khalid Hamandi<sup>4,5</sup> and Rhys H. Thomas<sup>1,2,3</sup>

<sup>1</sup>Department of Neurology, Royal Victoria Infirmary, Newcastle University, Newcastle-Upon-Tyne, United Kingdom

<sup>2</sup>Institute of Neuroscience, Newcastle University, Newcastle-Upon-Tyne, United Kingdom <sup>3</sup>Wellcome Trust Centre for Mitochondrial Research, Newcastle University, Newcastle-Upon-Tyne, United Kingdom <sup>4</sup>Cardiff University Brain Research Imaging Centre, Cardiff University, Cardiff, United Kingdom <sup>5</sup>Department of Neurology, University Hospital of Wales, Cardiff, United Kingdom

## 16.1 Introduction

The last 20 years have seen the field of epilepsy genetics move from detailed observation of clinical features in twins or families, to the discovery of classes of genes and types of variation associated with many of these epilepsies. The corollary of these discoveries has been the rapid introduction of clinically useful genetic testing in certain scenarios. It has been a surprise to researchers in the field just how useful some of these tests have become, creating the opportunity for precision medicine with targeted therapeutics, or avoidance of certain drugs in limited situations. Precision medicine remains an ambition in most cases, yet the prospect is only possible because of our better understanding of the genetic architecture of the epilepsies [1].

This chapter focuses on the role of clinical genetics in epilepsy. Epilepsy is a condition in which the brain has an enduring increased predisposition to seizures [2]. Seizures are a symptom of many clinically heterogeneous neurological diseases, so we shall focus on the *epilepsies*—those disorders in which seizures are the predominant manifestation, with or without the presence of other neurodevelopmental or neurodegenerative disorders. We will first summarize the nature of seizures and epilepsy in sufficient detail to facilitate shared understanding of phenotype between the geneticist, neurologist, and importantly the patient, second, review the role of genetic factors in epilepsy and its treatment, and finally, address the role of genetic testing in particular clinical scenarios.

## 16.2 What are seizures?

A seizure is a transient occurrence of symptoms and/or signs due to abnormal excessive or synchronous neuronal activity in the brain [2]. This definition emphasizes the necessity for a clinical manifestation of the underlying abnormal electrical activity, which may be recorded from the brain using electroencephalography (EEG). It should be recognized both that seizures can occur without being detected on scalp EEG and that epileptiform discharges can be detected on EEG without any clinical manifestation: while these require careful interpretation, they are not usually considered seizures. Diagnosis of a seizure can be difficult because in most cases the clinician does not observe the event in question—by their nature, seizures are often infrequent and rarely predictable. Usually the diagnosis of a seizure is made based upon the patient's or witness' description, increasingly frequently augmented by mobile phone video captured in the community.

The clinician seeks to compare the semiology (study of the signs of the seizure) of the event in question with that of recognized seizure types [3] and their mimics, in the context of the clinical scenario and the patient's previous history of seizures or other paroxysmal events. For example—a patient may be “absent” in a

seizure when they have a transient and self-resolving blank spell. With no further classification, this event is “dialeptic”—meaning there was an impairment of awareness [4]. With witness and clinical information, this could be categorized as broadly an absence seizure or a focal seizure with impaired awareness. However, the context is important—for example, absence seizures rarely begin in middle or late age, where the majority of seizures are of focal onset—in addition to findings from EEG and brain imaging.

The language can be confusing—but it is important that we get this right. There is no granularity in the terms “grand mal” (commensurate with a convulsion) and “petit mal” (a catch all for everything that is not a convulsion), and patients may see things differently to this using “grand mal” to describe big attacks with involuntary movements, which may be one of several seizure types. Furthermore, terms such as “complex partial seizure” are no longer included in the new classification schemes. The International League Against Epilepsy (ILAE) Task Force on Classification has produced an excellent web resource with the University of Melbourne, providing examples of many seizure types, their mimics, and associations: [www.epilepsydiagnosis.org](http://www.epilepsydiagnosis.org) [5].

Once diagnosed, most individual seizures are readily classified according to three major features: whether the onset (the semiology of a seizure can evolve during the seizure itself) was *generalized* (bi-hemispheric) or *focal* (uni-hemispheric), whether the onset had prominent *motor* features or not, and whether *awareness* was *impaired* during any part of the seizure [3,6]. The classification further subdivides motor and nonmotor onsets allowing for detailed phenotypic analysis, and separates bilateral tonic-clonic seizures according to their onset [*generalized tonic-clonic seizures* (bi-hemispheric onset), *focal to bilateral tonic-clonic seizures* (focal-onset seizures that propagate bi-hemispherically to cause what were previously called *partial onset seizures with secondary generalization*), or *tonic-clonic seizure of unknown onset* (accepting that sufficient evidence to ascertain the onset may not be available)]. Capturing a typical event with video-EEG remains the gold standard (Fig. 16.1).

### 16.3 What is epilepsy?

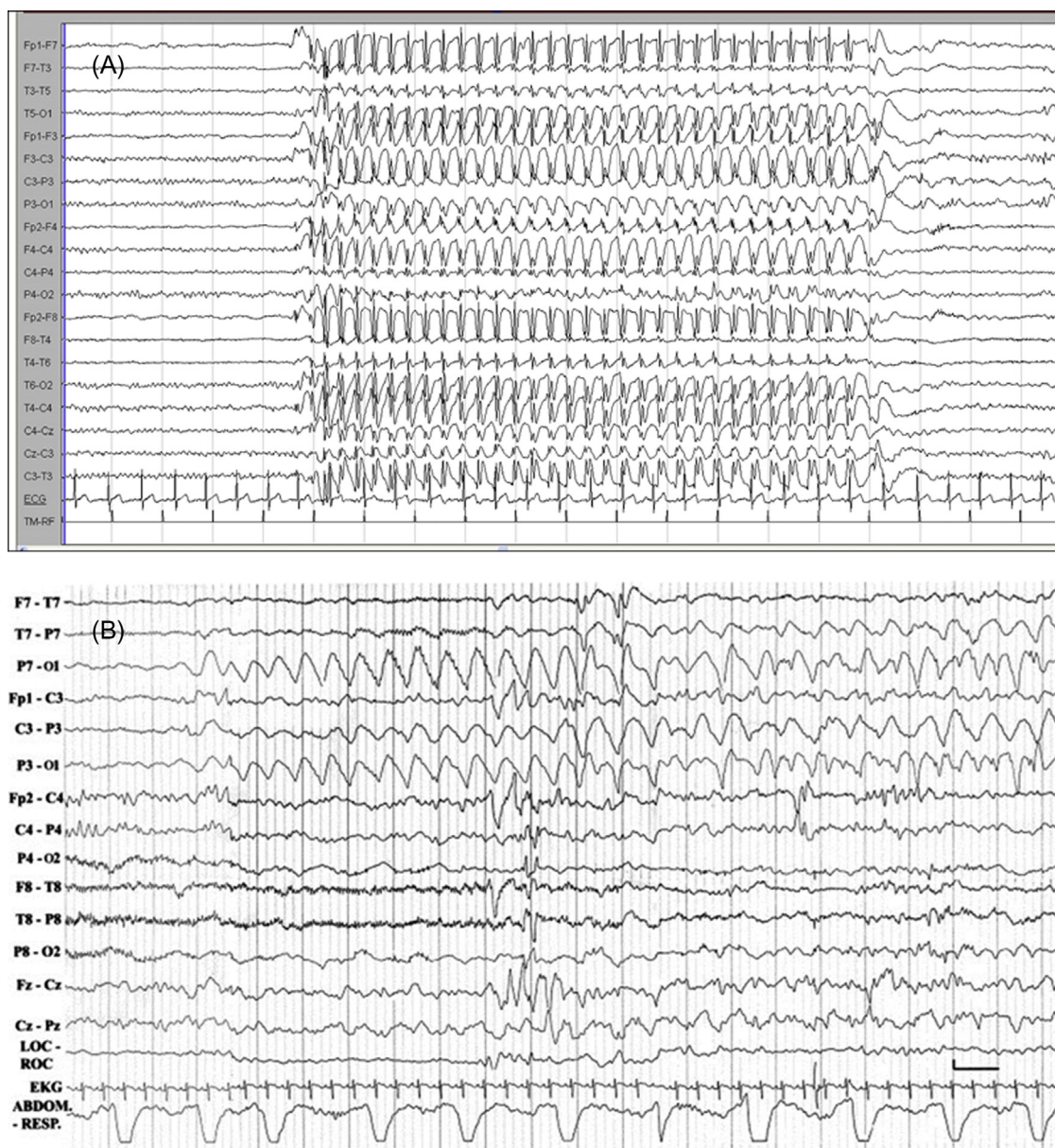
Epilepsy is diagnosed when a person’s brain has an enduring predisposition to generate epileptic seizures [2]. Epilepsy should not be diagnosed in everyone who suffers one, or even multiple seizures, because seizures can be provoked in a healthy brain at the time of an acute brain injury such as trauma, encephalitis, or stroke, or transient states of seizure predisposition such as alcohol withdrawal, electrolyte imbalance, hypoglycemia, or even fever in early childhood [8]. To diagnose epilepsy a clinician should be satisfied that the patient has had an unprovoked seizure and that their risk of further unprovoked seizures is >60% over the next 10 years [9], indicating the likelihood that the patient would benefit commencing antiepileptic treatment. The 60% threshold was selected because it approximates to the risk of further seizures after a second unprovoked seizure [10]. In reality, most patients referred to a clinical geneticist will have suffered multiple, often hundreds, of unprovoked seizures, clearly surpassing this diagnostic threshold. However, in some cases the patient may have suffered a single seizure has a constellation of clinical and EEG features that are recognized as a specific (*electroclinical*) *epilepsy syndrome* [5,11] or a demonstrable cause for an enduring seizure predisposition remains (e.g., an epileptogenic structural brain lesion).

The 2016 global prevalence of active epilepsy (seizure within previous 5 years) is estimated to be 0.6%, and half of these cases are thought to be idiopathic implying a predominantly genetic etiology rather than secondary to stroke, dementia, central nervous system infection, and other conditions that can result in epilepsy [12]. There is a bimodal prevalence (peaking in childhood and adolescence, before dropping during early-adulthood, and rising from mid-adulthood) [12]. The phenotypes of active epilepsy are very heterogeneous comprising of different seizure types, ages of onset and remission, seizure frequencies, treatment-responsiveness, and associated comorbidities. Sadly, antiepileptic drug treatment fails to achieve long-term seizure freedom in over a third of all people with epilepsy [13].

At the extreme end of the spectrum are the infantile-onset developmental and epileptic encephalopathies. Epileptic encephalopathies make up a heterogeneous group of syndromes with severe drug resistant epilepsy, often with many seizures per day and developmental delay, where the seizures themselves and abnormal electrical brain activity are thought to contribute to the developmental delay [14]. At the other end of the childhood epilepsies are those with later onset seizures, occurring once every few years and less overt comorbidities, and children with frequent but minimally intrusive seizures within self-limited syndromes such as childhood epilepsy with centrottemporal spikes (previously known as rolandic epilepsy). However, even this electroclinical syndrome is no longer considered to be “benign”—as there can be associated comorbid conditions, particularly learning disabilities—and antiepilepsy medication may be needed.

Idiopathic epilepsies alone (i.e., those without underlying structural or metabolic cause such as stroke, meningitis, encephalitis, trauma, or as yet occult causes) account for 182.6 disability-adjusted life-years per 100,000 population (0.56% of all cause disability-adjusted life-years) [12]. This is largely due to (1) seizures which can result





**FIGURE 16.1** (A) Ictal EEG during a typical absence seizure demonstrating bilateral spike-wave complexes at 3 Hz. This is a generalized seizure because both hemispheres are involved from the onset. (B) Ictal EEG during a focal seizure beginning in the left hemisphere (odd numbered leads) and spreading over to neighboring regions over the course of the visible recording. EEG, electroencephalography. Source: (A,B) Reprinted from Cerminara C, Coniglio A, El-Malhany N, Casarelli L, Curatolo P. Two epileptic syndromes, one brain: childhood absence epilepsy and benign childhood epilepsy with centrotemporal spikes, *Seizure* 2012;21(1):70–4. ©2012, with permission from Elsevier. (B) Reprinted from Patrizi S, Holmes GL, Orzalesi M, Allemand F. Neonatal seizures: characteristics of EEG ictal activity in preterm and fullterm infants. *Brain Dev* 2003;25:427–37, © 2003, with permission from Elsevier [7].

in delayed childhood development, loss of independence and work, loss of a driving license, falls and injuries, medication side-effects, emergency medical treatments, and premature death, (2) associated neurodevelopmental comorbidities, and (3) stigmatization.

## 16.4 Evidence for the genetic basis of epilepsies

The ILAE recognizes the increasing evidence and hence importance of genetic causes in common epilepsies [14]; to the point that despite the specific genetic basis being currently presumed rather than proven for the majority of individual cases, the evidence was considered sufficient for the *idiopathic generalized epilepsies* to be renamed the

**TABLE 16.1** Standardized incidence ratios (SIR) for the broad types of epilepsy in the relatives of probands with each broad type of epilepsy within the Genetic Epidemiology of Seizure Disorders in Rochester.

Epilepsy type in probands	Total relatives	Epilepsy type in relatives	With epilepsy	SIR (95% CI)
Generalized	708	Generalized	14	8.3 (2.93–15.31)
		Focal	8	2.5 (0.92 – 4.00)
		Unclassifiable	10	7.3 (2.82 – 12.98)
Focal	1239	Generalized	3	1.0 (0.00–2.19)
		Focal	16	2.6 (1.19 – 4.26)
		Unclassifiable	6	2.0 (0.40–3.85)
Unclassifiable	461	Generalized	6	5.5 (1.78–10.24)
		Focal	8	3.9 (1.39–6.46)
		Unclassifiable	3	3.5 (0.00–7.86)

Taken from Peljto AL, Barker-Cummings C, Vasoli VM, et al. Familial risk of epilepsy: a population-based study. *Brain* 2014;137(Pt 3):795–805 by permission of Oxford University Press not yet requested £102.08.

genetic generalized epilepsies [15]. Since William Lennox's pioneering studies through the mid-20th century (reevaluated using more modern definitions by Vadlamudi et al. [16]), twin studies have repeatedly demonstrated that epilepsy is highly heritable [16–19]. In the most recent analysis including 558 twin pairs the case-wise concordance of any seizure type in monozygotic twins was 0.64 compared to 0.19 in dizygotic twins [19]. There was also strong evidence that genetic factors are important in influencing epilepsy type with concordances of 0.79 in monozygotic twins versus 0.29 in dizygotic twins for generalized epilepsies (those characterized by generalized-onset seizures), 0.40 versus 0.03 for focal epilepsies, and 0.85 versus 0.25 for the genetic epilepsy with febrile seizures plus spectrum of syndromes (often associated with dominantly transmitted pathological variants in *SCN1A* [20]) [19]. While febrile seizures alone occurring between the ages of 3 months and 6 years are not indicative of epilepsy (fever may provoke seizures at this stage of development in many children with a healthy brain), these too show a strong genetic predisposition with concordances of 0.59 in monozygotic twins and 0.10 in dizygotic twins [19].

The risk of epilepsy in first-degree relatives of people with epilepsy is increased and recently quantified in a population-based study of 660 people with epilepsy in the Rochester Epidemiological Project and their 2439 first-degree relatives [21]. Among relatives of people with epilepsy the cumulative incidence of epilepsy up until 40 years of age was 4.7% (risk increased by 3.3-fold compared to the general population), with risk varying across different epilepsy types. The increased risk was greatest in the relatives of people with generalized epilepsies (6.0-fold) and of people with early-onset epilepsy and intellectual or motor disabilities (4.7-fold). In particular the incidence of generalized epilepsy in the relatives of a person with generalized epilepsy was greatly increased (8.3-fold). Conversely, as seen in the twin studies, focal epilepsies appeared less heritable and the risk of focal epilepsy in the relatives of a proband may have been similar regardless of whether the proband had a focal or a generalized epilepsy (Table 16.1). Interestingly, the transmission of epilepsy risk was greater from mothers (8.7-fold for offspring of mothers with generalized epilepsies and 4.0-fold for offspring of mothers with focal epilepsies) than from fathers. In fact, the prevalence of epilepsy in the offspring of fathers with epilepsy was not statistically different from that of the general population, although the point estimate of the standardized incidence ratio for the risk of epilepsy in the offspring of men with generalized epilepsies was 7.2-fold [21].

While phenotypes do vary within multiplex families with some families comprising some individuals with focal seizures, others with generalized seizures (and in some cases individuals whose epilepsy is characterized by both focal and generalized seizures), there is significant aggregation of epilepsy type within families [22,23] and of specific epilepsy syndromes and seizure types in both focal [23,24] and generalized epilepsies [23,25–28].

## 16.5 The genetic architecture of epilepsies

Together the epidemiological studies above suggest that a significant proportion of risk of a person developing epilepsy has a genetic basis. There appear to be genetic factors that are common to a broad epileptic

predisposition as well as others influencing specific features of the epilepsy phenotype, including seizure types and intellectual and motor comorbidities. The genetic architecture of common idiopathic epilepsies has been difficult to elucidate comprehensively [29].

While early genetic epidemiological studies suggested Mendelian or oligogenic inheritance of either complete syndromes or specific seizure or EEG traits, specific causes were rarely found in linkage or association studies, or if loci were found, they were rarely replicated, potentially because of significant phenotypic, locus, and allele heterogeneity [30]. A recurring theme has emerged in which genes harboring variants and copy number variant loci that have been implicated as the cause of epileptic encephalopathies or Mendelian familial epilepsies have subsequently been found to be implicated in sporadic cases of common epilepsies [31–35]. The following section summarizes what has been learnt about epilepsy genetics from genome-wide linkage and association studies, next-generation sequencing approaches, and karyotype and copy number variants associations, before describing a few recognizable epilepsies strongly associated to a single gene.

### 16.5.1 Linkage studies in families

The discovery of the association of the most important single gene with the epilepsies, that of *SCN1A*, encoding the alpha-1 voltage-gated sodium channel subunit (and which we will revisit throughout this chapter) resulted from linkage studies of two families with genetic epilepsy with febrile seizures plus phenotypes, one of the most common familial syndromes [36–38].

Linkage studies have also had some success in identifying Mendelian causes of rare and very phenotypically specific familial focal epilepsies such as autosomal dominant epilepsy with auditory features associated with *LGI* variants [39–41], and even significant success in epilepsies with notable locus heterogeneity such as autosomal dominant nocturnal frontal lobe epilepsy demonstrating associations with *CHRNA4* [42,43], *CHRNA2* [44], *CHRN2* [45,46], and *KCNT1*, [47] and genes implicated in a variety of epilepsy phenotypes such as and *DEPDC5* [33,48–53].

However, inconsistent results (possibly due to similar locus heterogeneity) have hampered the search for common loci linked to more common focal epilepsies such as childhood epilepsy with centrotemporal spikes (11p13 [54] and 15q14 [55], neither coinciding with *GRIN2A*, the gene rarely found to harbor variants in this condition [56]) and mesial temporal lobe epilepsy (3q [57], 3q26 [58], 4q13.2–q21.3 [59], 7p21.3 [60], 12q22–q23.3 [61], and suggested digenic inheritance via loci 18qter and 1q25–q31 [62]).

The largest genome-wide linkage study (including a metaanalysis) of 379 families (982 individuals) with genetic generalized epilepsies found no locus to be linked at the genome-wide significance level to genetic generalized epilepsy as a whole, although six loci [1p36.22, 3p14.2, 5q34 (interestingly close to the *GABA<sub>A</sub>* gene cluster), 13q12.12, 13q31.3, and 19q13.42] reached the suggestive threshold for linkage, which was more than expected by chance [63]. Analysis based on narrower phenotypes characterized by myoclonic and absence seizures found these to be significantly linked to 2q34 and 13q31.3, respectively [63]. Smaller but detailed studies of families with generalized epilepsy phenotypes have led to the association of generalized epilepsy to variants in *ME2* [64–66] and juvenile myoclonic epilepsy (or at least epilepsy with myoclonic seizures) being linked to chromosome 6 [67] and attributed to variants in *EFHC1* [68] and *ICK* [69], and to methylation of the promoter of *BRD2* [70]. However, these associations have not been supported by some larger independent studies [71–74]; hence, the role these genes should remain under critical review and may be population dependent [75,76].

### 16.5.2 Genome-wide association studies

Very much like the linkage studies, early GWAS of common variants with common epilepsies had mixed success. The only recurrent theme was that *SCN1A*—the predominant cause of Dravet syndrome (a rare form of childhood epilepsy previously known as severe myoclonic epilepsy of infancy) and of GEFS+ (a spectrum of seizure disorders spanning from isolated febrile seizures to severe epilepsies such as Dravet syndrome or epilepsy with myoclonic-atonic seizures previously known as Doose syndrome)—appears as a risk allele in many other forms of epilepsy.

A GWAS of broadly defined focal epilepsy found no association [77]. However, a subsequent study analyzing more phenotypically specific subsets found an association between mesial temporal lobe epilepsy with hippocampal sclerosis and a history of febrile seizures with a risk allele at 2q24.3 (around the *SCN1A* gene) [78]. The ILAE Consortium on Complex Epilepsies performed a metaanalysis and found two loci associated with epilepsy



as a whole (2q24.3—again implicating *SCN1A*—and 4p15.1 implicating *PCDH7*) and another associated with genetic generalized epilepsy (2p16.1) [79]. In 2018 the ILAE Consortium on Complex Epilepsies conducted a further “mega-analysis” of 15,212 people with a range of epilepsies and 29,677 controls [80]. Sixteen significant loci were identified near candidate ion channel, transcription factor, and pyridoxine metabolism genes. Three loci were found to associate with epilepsy as a whole: 16p12.1 and the previously demonstrated associations with 2p16.1 and 2q24.3 (the latter containing two independent signals). Focal epilepsy was associated with only one locus, at 2q24.3. The generalized epilepsies combined associated with 11 loci, with juvenile myoclonic epilepsy specifically associated with 16p11.2 and childhood absence epilepsy with 2p16.1 and 2q22.3. The heritability of genetic generalized epilepsy and nonacquired focal epilepsy attributable to common variants was estimated at 32.1% and 9.2%, respectively [80]. These findings confirmed the predictions that common variants would make a significant contribution to the risk of epilepsy and that focal and generalized epilepsies would be associated with both shared and distinct variants [81].

### 16.5.3 Rare variants in epileptic encephalopathies

The genetic epileptic encephalopathies are severe epilepsies usually manifesting with frequent seizures in infancy or early childhood and in which epileptic activity is thought to impede normal development or even cause regression [82]. There is significant phenotypic and genetic overlap with conditions causing epilepsy and intellectual disability because of the frequent difficulty in disentangling the relationship between neurodevelopmental trajectory and seizures, particularly in research using retrospective cases. There are many potential phenotypes, but only a minority can be categorized into specific electroclinical syndromes such as Ohtahara syndrome, early myoclonic epilepsy, epilepsy of infancy with migrating focal seizures, West syndrome, Dravet syndrome, epilepsy with myoclonic-atonic seizures, and Lennox–Gastaut syndrome among others. In many cases the phenotypes cannot be differentiated [83].

The giant stride in the understanding of these diseases came with the discovery through trio high-throughput sequencing studies that patients carry variants arising *de novo* in a large number of genes, many of which were already implicated in neuropsychiatric disorders [84–90]. Now, when modern clinical genetic tools such as karyotype, chromosomal microarray, epilepsy gene panel, mitochondrial DNA screen, and clinical whole exome sequencing are available around a quarter of children with early-onset epilepsies without clinical distinguishing features can receive an early genetic diagnosis [83]. This allows screening of parental DNA for the benefit of reproductive counseling based upon whether the responsible variant(s) occurred *de novo* [84] or was transmitted via the X chromosome (*PCDH19*-related epilepsies: a particularly interesting case because of its female-limited distribution—epileptogenesis is thought to result from the interaction of two populations of cells, each expressing a different allele as a result of lyonization) [91], autosomal recessively [92], or because of parental mosaicism [93]. Increasing diagnosis rates are facilitating genetically stratified follow-up studies of the natural history and treatment responses of each disorder that are increasingly guiding prognostication and treatment decisions [88,94,95]. However, the epileptic encephalopathies may not be purely monogenic diseases, and evidence is emerging for an oligogenic contribution to risk [89]. Studies over the coming years will hopefully identify the role of variation in these “modifier” genes that contribute to phenotypic variation, helping with not only prognostication but also the identification of candidate mechanisms for precision treatments.

### 16.5.4 Rare coding sequence variants in common epilepsies

Early studies sequencing ion channel genes or entire exomes of small numbers of people with genetic generalized epilepsies found a great deal of genetic variation at the individual level, failing to find statistically significant single variants, but identifying the need to use large cohorts analyzed using gene-based rather than variant-based association approaches [96,97]. The advent of high-throughput sequencing studies has provided evidence that rare (minor allele frequency <0.5%) and ultrarare (sufficiently rare to rarely recur and to be absent from public databases of population variation) sequence variants in genes implicated in rare monogenic epileptic encephalopathies and rare monogenic familial epilepsies are likely to be important in common idiopathic epilepsies [31–33,98].

The Epi4K study of ultrarare variants looked for enrichment of variants that were both predicted to have damaging consequences and which were absent from public variant databases within three cohorts (familial but

otherwise common forms of idiopathic generalized and focal epilepsies, and sporadic idiopathic focal epilepsies) compared to controls [31]. A set of genes comprising 43 genes known to cause dominant monogenic epilepsies was found to contain more variants of interest in people with familial idiopathic epilepsies (regardless of whether these were generalized or focal) than controls. However, this was not the case for the cohort with sporadic idiopathic focal epilepsies. Only the association of *DEPDC5* (classically associated with familial epilepsy with variable foci and familial forms of temporal lobe and frontal lobe epilepsy) with familial focal epilepsies reached study-wide statistical analysis in the primary collapsing analysis of all single genes (odds ratio = 8.1, 95% confidence interval (CI) 3.6–18.3) [31]. However, the genes ranked second to fifth in the nonacquired familial focal epilepsy cohort also have established roles in monogenic epilepsies: *LGII* (associated with autosomal dominant epilepsy with auditory features), *PCDH19* (the cause of a range of female-restricted X-linked epileptic encephalopathies spanning Ohtahara syndrome to a syndrome said to be comparable to Dravet syndrome), *SCN1A* (the principal cause of Dravet syndrome and genetic epilepsy with febrile seizures plus), and *GRIN2A* (associated with Landau Kleffner syndrome, the severest form of the childhood epilepsy with centrotemporal spikes spectrum). While no single gene reached formal study-wide significance in familial generalized epilepsies, a similar trend was seen to that in familial focal epilepsies: *KCNQ2* (a cause of self-limited familial neonatal seizures and also infantile epileptic encephalopathy), *GABRG2* (a cause of infantile epileptic encephalopathy and genetic epilepsy with febrile seizures plus), and *SCN1A* were in the top 10 ranks by enrichment.

Recently a study of people with familial and sporadic common generalized epilepsies found an enrichment of rare missense variants (minor allele frequency <0.5%) in a set 19 GABA<sub>A</sub> receptor subunit genes (odds ratio = 2.40, 95% CI 1.41–4.10), including *GABRG2*, and went on to validate and then to replicate the finding in independent data (odds ratio = 1.46, 95% CI 1.02–2.08) [32].

Building on the Epi4K study's findings in *DEPDC5*, a study of 121 people with sporadic idiopathic focal epilepsies found an increased burden of ultrarare variation in the 10 genes involved in mTOR signaling (particularly *DEPDC5*) even though (beyond focal seizures) the cohort had no evidence of tuberous sclerosis or focal cortical dysplasia [33].

A more phenotypically specific study of 194 sporadic cases of typical and atypical childhood epilepsy with centrotemporal spikes (the most common focal epilepsy of childhood) found an excess of rare, predicted damaging variants in *GRIN2A* (already associated with more severe epilepsies of the same spectrum) but no other gene or gene set [98].

Thus the genetic architecture of common epilepsies appears to include both common and rare sequence variants in genes whose roles have already been established in monogenic epilepsies.

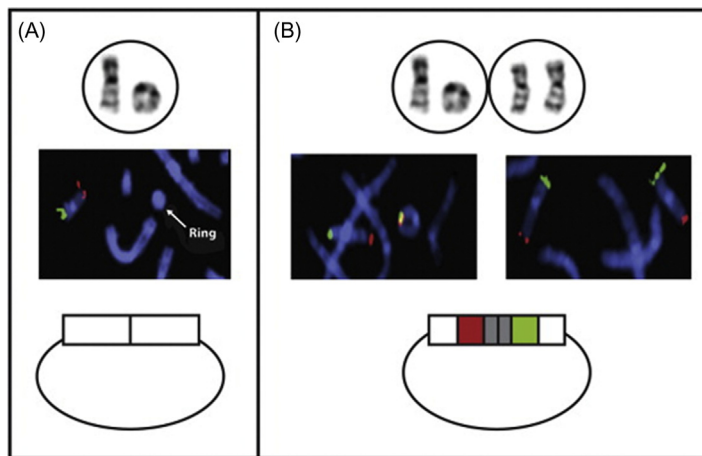
### 16.5.5 Noncoding variants

Noncoding variants were first implicated in epilepsy when Unverricht–Lundborg disease (progressive myoclonic epilepsy type 1A) was found to most commonly to be caused by expansions of a dodecanucleotide repeat in the promoter region of the *CSTB* gene, and in current clinical practice, exclusion of this remains the main role of screening for intronic variants [99].

Periventricular nodular heterotopia as a neuronal migration disorder exists in x-linked dominant, autosomal dominant, and recessive forms. Exonic variants in *FLNA* have been established as the principle cause of the x-linked form, which is usually lethal in the hemizygous state, often present with focal epilepsy in juvenile women associated with cardiovascular and musculoskeletal complications and sometimes borderline intellectual disability but sometimes cause an early-onset epileptic encephalopathy [100]. Somatic mosaicism of a *FLNA* intronic acceptor splice site [101], and more recently an intronic variant that resulted in abnormal inclusion of a poison exon usually skipped in neuronal progenitor cells [102], has been found to result in milder forms of x-linked periventricular nodular heterotopia, particularly in male cases.

Targeted sequencing study of intronic regions of *SCN1A* in individuals with phenotypes suggestive of Dravet syndrome who did not have pathological exonic variants in *SCN1A* found five intron 20 variants in a cohort of 640 individuals with developmental and epileptic encephalopathy unexplained by exon sequencing [103]. These variants resulted in the aberrant inclusion of a poison exon that led to reduced amounts of wild-type transcript in a cell model and two segregated with family members who had genetic epilepsy with febrile seizures plus phenotypes [103].

Whole-genome sequencing techniques are just beginning to reveal further roles of intronic variants in other epilepsies. The finding that familial adult myoclonic epilepsy (a rare autosomal dominant epilepsy) is caused by



**FIGURE 16.2** (A) Nonmosaic ring sample with no normal cells present. FISH with subtelomere probes (green = 20p, red = 20q) demonstrates that this nonmosaic ring has subtelomere deletions at one or both arms. (B) Mosaic ring sample with subtelomere repeats intact. FISH, Fluorescence in situ hybridization. Source: Reprinted from Daber RD, Conlin LK, Leonard LD, et al. Ring chromosome 20. *Eur J Med Genet* 2012;55(5):381–7, ©(2012), with permission from Elsevier.

a pentanucleotide repeat expansion [104] has raised the possibility that repeat expansions may play an underrecognized role in epilepsy, particularly those forms anecdotally suspected to show anticipation in some families.

### 16.5.6 Karyotypic abnormalities

A number of chromosomal abnormalities can cause epilepsies, but most are associated with dysmorphic or other striking extraneurological features. We shall focus on others such as inverted-duplication of chromosome 15 (InvDup(15)) and ring chromosome 20 (R(20)) as rare but well-characterized causes of epilepsy with variable intellectual disability and with little and variable or no typical dysmorphic features.

InvDup(15) is associated with epilepsy in two-thirds of the cases, often starting in childhood—a third during infancy with epileptic spasms [105]. A wide range of seizures is encountered including generalized tonic-clonic, focal, tonic, myoclonic, atonic, and atypical absence seizures, and status epilepticus and sudden unexpected death in epilepsy are common [105,106]. Intellectual delay is usual and at least moderate in most cases, and autism is common; thus InvDup(15) is a possible cause of Lennox–Gastaut syndrome (most typically due to tuberous sclerosis) [106].

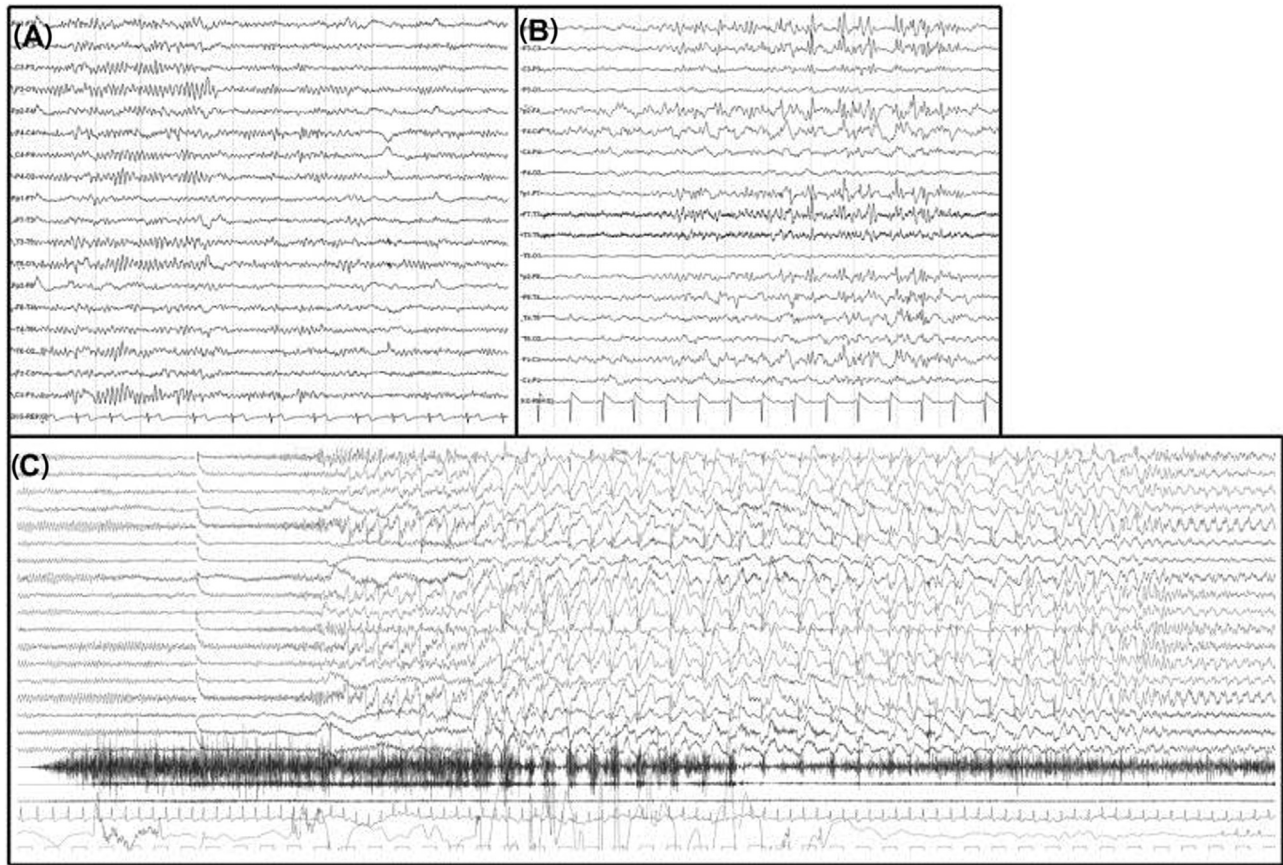
R(20) can be divided into two syndromes: the first has a more severe phenotype and resulting from fusion of the two arms of chromosome 20 during meiosis and resulting in significant subtelomeric deletions; the second is usually milder with later onset and results from mosaicism for a R(20) with less dramatic deletions (Fig. 16.2) [107]. Intriguingly, while the nonmosaic, a more severe form of R(20), has only ever been found to have occurred sporadically, the mosaic condition can be transmitted maternally due to transmission of an unstable chromosome 20 either in its linear shape but prone to ring re-formation or as a ring that in some cell lineages reopens [108].

Seizures typically begin in the first decade, and there may be mild or moderate associated learning difficulty, which can decline after onset of epilepsy. The majority of seizures are sleep-related with a focal, frontal onset semiology, and can be difficult to control. Generalized tonic-clonic seizures also occur. The electroencephalogram is relatively unique in the context of this syndrome, with runs of high-amplitude rhythmic frontal delta activity that last minutes to hours even when seizures are not taking place, or as nonconvulsive status epilepticus during which the patient may appear confused (Fig. 16.3). The mechanism causing seizures is yet to be fully elucidated but is likely to include the nicotinic acetylcholine receptor alpha-4 subunit (*CHRNA4*, associated with autosomal dominant nocturnal frontal lobe epilepsy) and voltage-gated potassium channel subfamily KQT type 2 (*KCNQ2*, associated with self-limited familial neonatal seizures and also the severe epileptic encephalopathy Ohtahara syndrome), which may be lost within the terminal deletion of the long arm of chromosome 20. R(20) is probably underdiagnosed and should be considered in anyone with drug-resistant frontal lobe seizures from sleep, particularly in the context of the typical electroencephalogram.

### 16.5.7 Copy number variation

Copy number variants (CNVs), and particularly microdeletions, are important contributors to the risk of pediatric-onset epilepsies and neurodevelopmental disorders. Despite being widely called “pathogenic”—the





**FIGURE 16.3** (A) Normal waking EEG. (B) Typical finding of bi-frontal sharp waves in the interictal EEG of a r(20) patient. (C) Ictal EEG recording during nonconvulsive status epilepticus characterized by diffuse low-voltage activity following by recruiting rhythm over the frontal areas. After 15–20 s, there is appearance of discharges of sharp and slow waves with anterior predominance. Clinically, the patient showed motor ictal semiology (tonic contraction of both arms followed by a clonic phase with right predominance). EEG, Electroencephalography. Source: Reprinted from Daber RD, Conlin LK, Leonard LD, et al. Ring chromosome 20. *Eur J Med Genet* 2012;55(5):381–7, ©(2012), with permission from Elsevier.

recurrent CNVs are neither necessary nor sufficient to cause epilepsy and are present in people with epilepsy as well as controls. The phenotypic heterogeneity of recurrent CNVs that span neurodevelopmental disorders and near-normal intellectual functioning provides a daunting clinical challenge for the geneticist when it comes to providing accurate familial counseling about risk [109]. Recognition and prediction of when CNVs may be involved is difficult but is typically included in addition to exome sequencing as a typical strategy for investigating early-onset epilepsies [110]. Analysis of CNVs has added depth but not yet understanding to the complex genetic architecture of epilepsy as a whole. It is blurring the lines between monogenic and polygenic background where it is evolving as a contributor to causation across a wide range of epilepsies [111]. It adds detail to previously characterized but heterogeneous syndromes such as glucose transporter 1 (GLUT1) deficiency syndrome [112] and CNVs can potentially influence epilepsy pharmacogenomics [113].

Three recurrent microdeletions (15q11.2, 15q13.3, and 16p13.11) and rare (private) CNVs are more common in individuals with genetic generalized epilepsies [114–117], particularly affecting genes already implicated in neurodevelopmental disorders [118–122]. Analysis of large cohorts with heterogeneous epilepsies has demonstrated that in particular, the 16p13.11 deletion is associated with a diverse phenotypic spectrum [123], and rare CNVs are more common if there is an associated intellectual disability [124,125]. Large CNVs are frequently found in patients with epileptic encephalopathies such as infantile spasms and Lennox–Gastaut syndrome [126,127]. Nonlesional focal epilepsies are more rarely associated with CNVs [116,117,122]. Rarer phenotypes, atypical variants of common epilepsies, and epilepsies with structural abnormalities on imaging or dysmorphism also seem to

be associated with CNVs [128]. These do not appear sufficient to cause epilepsy but appear to contribute to a general susceptibility to neurodevelopmental disorders. Recently, rare and small (<50 bp) microdeletions and microdeletions within coding and noncoding regions within 100 kbp of known epilepsy genes have been found to be enriched among people with a range of epilepsies [129].

## 16.6 Mitochondrial epilepsies

Mitochondria are ubiquitous, and vital cellular organelles are present in all cells apart from erythrocytes. They are involved in multiple cellular processes, including apoptosis, senescence, and biochemical homeostasis; however, their principle function is generation of energy from metabolic pathways, via oxidative phosphorylation. In practice, most of the recognized mitochondrial disorders are caused by dysfunction of the respiratory chain and the consequent defective energy metabolism. Consequently, mitochondrial disease often manifests most prominently in energy-hungry systems such as skeletal muscle and the central nervous system.

Mitochondria are unique in that they possess their own genetic code, separate to the somatic nuclear DNA. The mitochondrial genome is 16.6 kbp, with a noncoding region, the displacement loop, controlling the replication of the mtDNA. It codes for 13 peptides that are subunits of the oxidative phosphorylation system, and 24 RNA molecules that code for other proteins within the mitochondria—the rest of the 1500 different proteins needed for the mitochondria to function properly are derived from nuclear DNA. Interestingly, despite this disparity in contribution, 80% of adult onset mitochondrial disease is due to a defect in the mitochondrial genome. This can provide a valuable clue to diagnosis, as mitochondrial DNA is inherited matrilineally. In children, nuclear gene mutations account for the majority of mitochondrial disease (75%–80%), with autosomal recessive inheritance being most common.

### 16.6.1 Heteroplasmy

There are thousands of mitochondria per cell, not only each carrying their own genetic sequence, but there can be mitochondria with different genetic codes within each cell—the phenomenon of heteroplasmy. This heteroplasmy can vary between cells, between tissues and even with age in the same individual, and the threshold at which the mutated: wild-type mitochondrial DNA ratio causes disease, is also variable. This creates even more clinical diversity within an already heterogenous group of patients. Heteroplasmy therefore increases the challenge of diagnosing mitochondrial conditions, as the pathological variant may be minimally expressed in leukocytes but present at different levels in other tissues.

### 16.6.2 Clinical phenotypes

Due to the omnipresence of mitochondria, epilepsy due to a mitochondrial disorder may be accompanied by a range of features in other organ systems. Certain mitochondrial disorders may present as a recognized syndrome, but some presentations are vague and “nonsyndromic.” An enhanced level of suspicion and honed clinical acumen is needed to diagnose the mitochondrial dysfunction in these cases. A clinical picture which seems to involve multiple organ systems should prompt suspicion, especially when accompanied by a suggestion of impaired energy metabolism. Biomarkers that point to disruption in energy metabolism, such as raised lactate, fatty acid or organic acid metabolites or Krebs cycle enzymes, are useful when considering whether a patient may have a mitochondrial disorder. Functional studies demonstrating respiratory chain deficiencies, particularly complex 1 deficiencies, provide stronger supportive evidence. Genetic testing is progressively improving the diagnostic accuracy of these conditions.

### 16.6.3 Recognized mitochondrial epilepsy syndromes

Mitochondrial disorders where the predominant phenotype includes prominent seizures include MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes), Leigh syndrome, Alpers–Huttenlocher syndrome and MEMSA (myoclonic epilepsy, myopathy and sensory ataxia) in childhood onset cases, and MELAS, MERRF (myoclonic epilepsy with ragged red fibers), and NARP (neuropathy, ataxia, and retinitis pigmentosa) in adult onset cases [130].

MELAS is the archetypical mitochondrial disorder. It is usually caused by a mutation in the MT-TL1 gene, most commonly m.3243A > G. There are multiple epilepsies here: the stroke-like episodes that characterize this condition are thought to be related to acute seizures; progressive atrophy and acute damage can predispose to focal seizures; and myoclonic seizures are also seen. The progression of the encephalopathy may be slowed by proactive use of anticonvulsants. In Leigh syndrome, there is psychomotor regression, which can be episodic and sometimes rapidly and relentlessly progressive, but sometimes taking a more indolent course. MRI will show lucent lesions in the basal ganglia and brainstem (and the mimic, biotin- or thiamine-responsive encephalopathy should be excluded) [131]. Epilepsy is often a feature with a wide variety in the pattern of seizure types and variety in the response to treatment. Often, multiple seizure types will coexist. There are over 75 genes, nuclear and mitochondrial, currently recognized as causing Leigh syndrome. The epilepsy in Alpers–Huttenlocher is intractable and forms a triad with liver disease and psychomotor regression. The condition is due to mutations in the nuclear *POLG* gene—one of many manifestations of this particular gene.

MEMSA (previously known as spinocerebellar ataxia with epilepsy) is a disorder most commonly caused by recessive mutations in *POLG*. The eponymous myoclonic epilepsy in this syndrome may be accompanied by other seizure types and nonepileptic myoclonus. Interestingly, these seizures can be very stereotyped, with focal onset in a limb, and then with secondary generalization. The seizures are typically a later sign, preceded in late childhood/young adulthood by cerebellar ataxia, followed by progressive proximal myopathy and frequently encephalopathy. Valproate is absolutely contraindicated in those with pathogenic *POLG* mutations as it can precipitate fulminant liver failure; it should be used with extreme caution in mitochondrial disease in general. Caution should be taken when introducing certain antiepilepsy drugs in case they precipitate lactic acidosis.

MERRF has a variable age of onset from childhood to mid-to-late adulthood, and most frequently (over 80%) the cause is m.8344A > G. Common features are a generalized epilepsy, ataxia, weakness, and, over time, an acquired dementia. Other associated features include hearing loss, short stature, optic atrophy, and cardiomyopathy with Wolff–Parkinson–White syndrome, pigmentary retinopathy, and lipomatosis.

## 16.7 Progressive myoclonic epilepsies

The progressive myoclonic epilepsies are a collection of monogenic diseases by the onset of myoclonic seizures (jerks that range from subtle to violent) in previously healthy children or adolescents, which in many cases progress to become very disabling within a few years, particularly in those patients whose seizures are photosensitive and or worsened by manual activity (praxis-induced). The initial differential diagnosis of a progressive myoclonic epilepsy and even its recognition in a patient thought to have a severe juvenile myoclonic epilepsy can be difficult, particularly in the majority who have no positive family history. However, individual disease-specific geographic or ethnic associations (East Asia, North Sea, Finland, Ashkenazi), the presence of a family history with anticipation in dentatorubral–pallidoluysian atrophy, and phenotypic associations (other seizure types, ataxia and other movement disorders, visual loss, visual hallucinations, dementia, and extraneurological manifestations such as renal failure, hepatosplenomegaly, short stature, or dramatic lipomata), test results (giant somatosensory-evoked potentials, or of Lafora bodies on axillary skin biopsy), and prognostic course can help [132]. They often require specific thought with regards to testing because most are autosomal recessive [Unverricht–Lundborg disease (*CSTB*), Lafora body disease (two genes: *EPM2A* and *NHLRC1*), the neuronal ceroid lipucinoses, sialidosis type 1, and the GM2 gangliosidoses], dentatorubral–pallidoluysian atrophy results from trinucleotide expansion in *ATN1* and Unverricht–Lundborg disease from intronic dodecanucleotide repeats in *CSTB*, and MERRF requires screening for mitochondrial DNA mutations.

## 16.8 Pharmacogenetics of epilepsy

### 16.8.1 Human leukocyte antigens and adverse antiepileptic drug reactions

A great deal of research on the genetic risk factors for type IV hypersensitivity reactions to antiepileptic drugs has been performed across a range of ethnicities. Most reactions manifest dermatologically as a mild maculopapular exanthema, but more severe cases can present as toxic epidermal necrolysis, Stevens–Johnson syndrome, and systemic hypersensitivity reactions such as drug-related eosinophilia and systemic symptoms. The most

established associations are between exposure to aromatic antiepileptic drugs (carbamazepine and its derivatives, phenytoin, and lamotrigine) and human leukocyte antigens (HLA): HLA-B\*15:02 (most commonly found in Han Chinese), HLA-A\*31:01 (Japanese and Korean), and HLA-A\*24:02 (Japanese, Korea, and Taiwanese) [133]. HLA genotyping is a highly predictive test that could stratify the risk of hypersensitivity, allowing enabling more personalized treatment [134]. The Food and Drug Administration recommends screening for HLA-B\*15:02 before commencing carbamazepine in those from or that have ethnicity from in Eastern Asian populations [135]. However, screening of this nature would not exclude other HLA genotypes that are related to serious skin reactions [136]. While there is some cross-reactivity to different drugs, this remains difficult to predict on an individual basis, and once a patient has suffered an idiosyncratic hypersensitivity reaction to one aromatic antiepileptic, they should only be challenged with others when necessary and with great caution.

### 16.8.2 Cytochrome enzymes and antiepileptic medication

Phenytoin is a potentially toxic antiepileptic drug largely metabolized by the liver. The first step is hydroxylation by cytochrome P450 (CYP) enzymes. CYP2C9 usually conducts 90% of this [137], but polymorphisms can cause significant changes [138]. Rarer CYP2C9 variants can be null alleles (e.g., \*6) leading to a very significant reduction in phenytoin clearance [139]. Before a prescription is issued, it would be informative to be aware of the genotype to enable avoidance of phenytoin or guidance of dose [140]. More recently, a GWAS of people with European ancestry and phenytoin-induced hypersensitivity identified genetic predictors in *CFHR4* and *CFH*, members of the complement factor H-related protein family [141]. However, at present, no authority requires genotyping before prescription of phenytoin despite potential cost savings in antiepileptic pharmacogenotyping [142].

### 16.8.3 Sodium channel genes and drug response

The response of patients to sodium channel blocking antiepileptic drugs is very variable. In many cases of focal epilepsies, these drugs are very effective and considered first line [143], and in many patients with Dravet syndrome (caused by variants in *SCN1A*) or myoclonic seizures, they can be unhelpful (and it is recommended that they are avoided), but this is not always the case. Studies of the association between common variants in sodium channel genes (particularly *SCN1A*) and responses to sodium channel blocking drugs have been far from unanimous [144–146]. In epilepsies caused by variants of large effect in *SCN2A*, an association has been found between the age of onset (which predicts gain-of-function vs loss-of-function consequences) and seizure responsiveness to sodium channel blocking drugs [147], but we have not yet reached the point of prospective studies to assess directed treatment.

### 16.8.4 Hyponatremia and sodium channel blocking antiepileptic drugs

Hyponatremia is a common side effect of carbamazepine and oxcarbazepine, and while it is usually mild and well tolerated, in some cases can be severe enough to require a change of treatment. No genomic association with this reaction has identified [148].

## 16.9 Molecular genetic testing strategies for epilepsy

Targeting of diagnostic genetic tests needs to be in harness with a neuropsychiatrician or adult neurologist with a special interest in epilepsy, and geneticist and related teams to design the optimum testing strategy. Furthermore, the “best” testing scheme will depend on the clinical scenario, question posed, and the resources available.

There are no international protocols to guide as to who should receive testing, and it is clear that we are underutilizing genetic testing in epilepsy. A rule of thumb in the adult clinic is that “epilepsy alone” is insufficient and that “epilepsy plus” is needed. This may be epilepsy and one or more of autism, intellectual disability, or dysmorphism; it is currently unclear whether a family history or unusual phenotype across generations is sufficient. In some situations, drug-refractory epilepsy or epilepsy with a family history may be sufficient to support a clinical testing approach. The highest yield is for early-onset epilepsy—and when there is “epilepsy plus” with seizures that start before the age of 6 months—providing that this is not a hypoxic ischemic injury clinically. The



yield of testing in children with epilepsy presenting in the first few years of life is high: 40% overall, much higher in the presence of features of tuberous sclerosis, metabolic disease, or brain malformations, but still around 25% in cases without distinguishing features [83].

### 16.9.1 Genetic testing methods

*A single gene test:* the needle in the haystack. There are few clinical scenarios in which this is recommended, typically in common and phenotypically striking disorders, such as tuberous sclerosis, or for genes poorly covered by high-throughput sequencing panels. Whereas it is profoundly possible to diagnose disorders with strong monogenic associations such as Dravet syndrome and GLUT1 deficiency syndrome (Box 16.1) on a clinical basis and to confirm this with a single test (see Box 16.1)—a negative report to a single gene test would then commence a cascade of testing—probably to include an exome panel. It is still probably better to start with a gene panel in most scenarios than to elect for a single gene test, and in the case of severe infant-onset epilepsies, early-targeted exome sequencing may be quicker and more economic than a staged approach to investigations [149].

*Comparative genomic hybridization (CGH) arrays* for clinically relevant copy number variation. This can be considered a screening test because it is cheap, swift to report, and yet unlikely to produce results that will result in a change to therapy. Unless the rearrangement is recurrent and a known genomic syndrome, the larger or more gene rich the CNV, the easier it is to interpret.

*Gene panels* of targeted sequencing of multiple genes to search for mutations with proven disease associations—with epilepsy (and typically neurodevelopmental disorders)—is currently the most cost-effective genetic testing strategy in epilepsy. Many providers are performing whole exome or whole genome and then reporting the variants in a preidentified number of genes. This has the advantage that reanalysis by requesting further “add on panels” are swift and less expensive.

There is a range of commercial and government-funded (in the United Kingdom through the National Health Service) providers. No two gene panels or providers seem alike, in terms of the number and range of genes tested [88]. Also gene panels for epilepsy typically include other neurological disorders where recurrent seizures are a prominent feature, but the clinical diagnosis may be more obvious from nonseizure features, such as dysmorphism, metabolic disturbances, or other clinical or paraclinical findings. In the context of clinical rather than research testing, there is a dramatic difference in the diagnostic yield across the many genes typically included in panels with *SCN1A*, *KCNQ2*, *CDKL5*, *SCN2A*, *PRRT2*, *PCDH19*, *STXBP1*, *SLC2A1*, *GABRG2*, *SCN8A*, *UBE3A*,

#### BOX 16.1

##### Glucose transporter1 deficiency syndrome

An example of the utility of thorough testing of a single gene is for the identification of a variant in *SLC2A1* in confirming a diagnosis of glucose transporter 1 (GLUT1) deficiency syndrome, a spectrum of clinical manifestations resulting from decreased glucose concentration in the central nervous system. Most cases appear to be autosomal dominant, but recessive forms have been reported, and while most variants are sequence variants, 11% of patients may have deletions of entire exons requiring exon-level array comparative genomic hybridization or multiplex ligation-dependent probe amplification to exclude the diagnosis [156]. Additionally splice-site and even deep intronic variants may be important in a minority [157]. A typical clinical history for the autosomal dominant disorder would be of an epilepsy beginning in early childhood (under the age of 3 years) with short blank events. These are characterized as “typical absence seizures” with supporting a supporting electroencephalographic signature of generalized spike wave activity. However, the onset of typical absence seizures at such an early age is unusual (childhood absence epilepsy usually presents between 4 and 10 years of age), and there may be an association of seizure occurrence to meal times. This could account for 10% of absence epilepsy depending on age, being particularly common in those with younger onset absence seizures. Later life manifestation can include movement disorders with paroxysmal dyskinesias, again related to meal times, in some cases leading to diagnostic confusion or a label of nonepileptic attack disorder. There can also be some intellectual disability. While epilepsy is not the only manifestation of GLUT1 deficiency syndrome, the ability to perform a well-described genetic test has largely superseded unreliable and invasive studies for fasting hypoglycorrhachia (low glucose in cerebral spinal fluid), and reduced misdiagnosis and misclassification in defining a distinct genetic epilepsy. Early diagnosis can lead to rapid prognostication based on functional assays [158] and more personalized therapies (classically, GLUT1 deficiency syndrome responds particularly well to a ketogenic diet rather than antiepileptic medication).

**BOX 16.2**

A list of genes suggested for inclusion in a gene panel for general epilepsy based on current published evidence of disease causation and yield. (This does not include screens for causes of progressive myoclonic or mitochondrial epilepsies.)

ALDH7A1, ALG13, AMT, ANKRD11, ARHGEF9, ARID1B, ARX, ASXL3, ATP1A2, BRAT1, CACNA1A, CACNA1E, CASK, CDKL5, CHD2, CHRNA4, CHRN2B, CNTNAP2, COL4A3BP, DCX, DDX3X, DEPDC5, DNM1, DOCK7, DYNC1H1, DYRK1A, EEF1A2, FOXG1, GABRA1, GABRB1, GABRB2, GABRB3, GABRG2, GLUD1, GNAO1, GNB1, GPHN, GRIN1, GRIN2A, GRIN2B, GRIN2D, HCN1, HNRNPU, HUWE1, IQSEC2, ITPA, KCNA2, KCNB1, KCNH1, KCNJ10, KCNMA1, KCNQ2, KCNQ3, KCNT1, NEXMIF, LGI1, MAPK10, MBD5, MECP2, MEF2C, MTOR, NACC1, NPRL2, NPRL3, NRXN1, PCDH19, PIGA, PLCB1, PNKP, POLG, PRRT2, PURA, QARS, RELN, RORB, RYR3, SCN1A, SCN1B, SCN2A, SCN8A, SIK1, SLC1A3, SLC2A1, SLC6A1, SLC6A8, SLC9A6, SLC12A5, SLC13A5, SLC25A22, SLC35A2, SMARCA2, SMC1A, SNAP25, SPTAN1, STX1B, STXBP1, SYN1, SYNGAP1, SYNJ1, TBC1D24, TCF4, TSC1, TSC2, UBE3A, WDR45, WWOX, ZEB2.

*MECP2*, *GRIN2A*, *TSC2*, *FOXG1* being most common in a review of samples by one clinical diagnostics service [150].

The tests for epilepsy genes can be divided broadly into the following categories:

1. Single genes causing epilepsy and/or epileptic encephalopathy (e.g., channelopathies)
2. Probably associated with epilepsy (e.g., genes associated with epilepsy in animal models only)
3. Multisystemic genetic syndromes that can also cause seizures (e.g., cardiofacialcutaneous syndrome)
4. Neurological syndromes (e.g., Joubert syndrome)
5. Metabolic disorders
6. Lysosomal disorders
7. Congenital disorders of glycosylation
8. Peroxisomal biogenesis disorders
9. Genes important in pharmacogenomics
10. Genes currently not known to cause epilepsy (e.g., Brugada syndrome)

Box 16.2 provides a suggested list of genes that could be included in a general epilepsy panel, with a particular focus on epilepsy with comorbid intellectual disability, or a history suggestive of epileptic encephalopathy—this is the context in which we most frequently undertake panel testing in the adult epilepsy clinic (Box 16.2).

*Mitochondrial testing* is not the logical next stage in a cascade of testing but should always be considered in the correct clinical context, suggesting a mitochondrial disorder, and when an exome/genome panel has been unrevealing.

*Karyotyping* retains an importance, specifically analysis for ring chromosomes without deletions in people with intellectual disability and clusters of seizures/recurrent status epilepticus. It should be extended sufficiently to exclude mosaicism. It is now replaced by array CGH for the detection of genomic deletion or duplication.

*Repeat expansion disorders* need to be considered in familial and individuals with progressive myoclonus epilepsies. The most common mutation that causes Unverricht–Lundborg disease is an unstable expansion of a 12 nucleotide, dodecamer repeat (5'-CCCCGCCCCGCG-3'), in the promoter region of the *CSTB* gene.

*Progressive myoclonic epilepsies* can now be screened for by gene panels, and these can be a good first step when there are no clues as to aid differential diagnosis. However, the interpretation of rare variants in recessive disorders and the exclusion specific forms can be challenging, so an approach supported by biochemical and microscopy investigations is suggested (Table 16.2).

### 16.9.2 Limitations to current genetic testing strategies

With strong evidence of heritability, efforts to find candidates to explain a majority of both focal and generalized epilepsy started with much optimism. For research the major limitation was access to sufficient numbers of well-phenotyped patient samples, clinical heterogeneity, the need for expert clinical assessment in defining phenotypes rather than diagnostic tests and ill-defined concepts added to this challenge. These barriers are gradually being overcome. One factor that remains is that the majority of patient samples that have been amassed have a white European background.



**TABLE 16.2** Biochemical, microscopic, and genetic investigations to consider in the diagnosis of progressive myoclonic epilepsies.

Disorder	Investigations
Unverricht–Lundborg disease (EPM1)	<i>CSTB</i> mutation analysis (repeat disorder)
Lafora body disease (EPM2)	Skin biopsy for Lafora bodies <i>EPM2A</i> and <i>NHLRC1</i> mutation analysis
Action myoclonus renal failure syndrome (EPM4)	<i>SCARB2</i> mutation analysis
<i>PRICKLE1</i> -related progressive myoclonus epilepsy with ataxia (EPM5)	<i>PRICKLE1</i> mutation analysis
“North Sea” progressive myoclonus epilepsy (EPM6)	<i>GOSR2</i> mutation analysis
MERRF	Plasma lactate and pyruvate (as screen) Muscle biopsy microscopy for ragged-red fibers <i>MT-TK</i> mutation analysis
NCLs	Skin biopsy electron microscopy for deposit profiles specific to each NCL Leukocyte enzyme analysis of PPT1, TPP1, and CTSD activity <i>PPT1</i> , <i>TPP1</i> , <i>CLN3</i> , <i>CLN5</i> , <i>CLN6</i> , <i>MFSD8</i> , <i>CLN8</i> , <i>CTSD</i> , <i>DNAJC5</i> , <i>CTSF</i> , <i>ATP13A2</i> , <i>GRN</i> , <i>KCTD7</i> mutation analysis
DRPLA	<i>DRPLA</i> mutation analysis (repeat disorder)
Sialidosis type I (cherry-red spot myoclonus syndrome)	Sialo-oligosaccharides in urine Leukocyte enzyme analysis of neuraminidase activity <i>NEU1</i> mutation analysis
Tay–Sachs disease (GM2 gangliosidosis)	Leukocyte enzyme analysis of hexosaminidase A and B activity <i>HEXA</i> mutation analysis
Gaucher disease	Leukocyte enzyme analysis of $\beta$ -glucocerebrosidase <i>GBA</i> mutation analysis

DRPLA, Dentatorubral–pallidoluysian atrophy; MERRF, myoclonic epilepsy with ragged red fibers; NCLs, neuronal ceroid lipofuscinoses.

Adapted from Malek N, Stewart W, Greene J. The progressive myoclonic epilepsies. *Pract Neurol* 2015;15(3):164–71. Permission from BMJ not yet requested £506.09.

Despite the attractions, genetic testing is not yet appropriate as the first-line testing for most adults with a common epilepsy. At present, there are near to 1000 putative epilepsy-associated genes [151], and blanket testing creates a vast number of variants of uncertain clinical significance. It is a challenge to decide which group of genes to include in certain patient populations in an efficient and cost-effective testing strategy. Thought should be given to staged testing such as karyotyping to examine for ring chromosomes where an exome panel has not provided an answer where there is a strong suspicion of a genetic epilepsy. Biparental trio exome studies facilitated by the next-generation sequencing revolution have demonstrated that up to 20% of epileptic encephalopathies may be attributable to de novo variation [88]—but also, occasionally, low-level mosaicism in a parent [93], this is now a challenge for the limits of our ability to detect mosaicism and counsel parents.

Even when genotyping suggests an abnormality such as in sodium channels, this does not predict pathogenicity as missense and truncation may result in heterogeneity. Relatively infrequently, but increasingly so, an etiological diagnosis results in specific treatment [152]. Whereas phenotypic severity may not be clear, such an identification can guide medications with an avoidance of sodium channel blockers in Dravet syndrome [153] or encourage the use of sodium channel blockers in patients with gain of function *SCN2A* encephalopathies [147]. There are traps in moving resources where this may be unnecessary or leads to poorly characterized results. Where a result suggests a variant of unknown relevance, there may be uncertainties how to confirm or refute a pathology. Invasive tests may be required with metabolic disorders. mTORopathies can give cortical malformations, but there may be no change evident on structural imaging [33].

Variants may contribute to both autosomal dominant and recessive inheritance. There is a missing heritability with only a fraction explained thus far [29], but new findings such as noncoding variants that affect splicing of *SCN1A* in Dravet syndrome [103] are in need of replication in other epilepsy types [111], and unexplained cases will need periodic reanalysis to increase yield [154].

Finally, the results of single research studies should be accepted with caution and validated in independent cohorts while considering the possibility of different mechanisms of disease in different ethnicities. The age of

large studies using whole exome and whole-genome sequencing will enable us to reappraise the role of genes previously implicated in smaller studies. It is likely that the pathogenic roles of some genes identified in even well-conducted studies may need to be revised in light of new evidence [155]. This will require substantial collaborative efforts to curate databases containing large numbers of validated and candidate genes (and disease-causing variants) implicated in epilepsies and guidance for their interpretation in specific clinical contexts.

## 16.10 Summary

Clinical management of a patient's epilepsy depends crucially on its classification, which guides prognostication and treatment strategies. In addition, clinical research to elucidate pathogenesis and identify opportunities to intervene involves the identification of variables that associate with and potentially explain the clustering of phenotypic features into discrete syndromes. In order to best facilitate advances in precision medicine, classifications should place individuals into biologically and clinically homogeneous groups that represent natural classes even when specific biological and clinical factors remain to be discovered. We have long moved "beyond the ion channel." The search for single genes for single phenotypes in epilepsy has been largely superseded as in many areas of neurology. Gene discoveries of monogenic variants of large effects in large families add to the body of knowledge, but for common epilepsies, the variants in a number of genes described above are likely contributory. The implicated genes have biological functions as diverse as coding for ion channel subunits, transcription factors, mTOR signaling, vitamin B-6 metabolism, and chromatin remodeling, thus demonstrating how wide the variation in pathways may be.

## References

- [1] Thomas RH, Berkovic SF. The hidden genetics of epilepsy—a clinically important new paradigm. *Nat Rev Neurol* 2014;10(5):283–92.
- [2] Fisher RS, van Emde Boas W, Blume W, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia* 2005;46(4):470–2.
- [3] Fisher RS, Cross JH, French JA, et al. Operational classification of seizure types by the International League Against Epilepsy: position paper of the ILAE commission for classification and terminology. *Epilepsia* 2017;58(4):522–30.
- [4] Luders H, Acharya J, Baumgartner C, et al. Semiological seizure classification. *Epilepsia*. 1998;39(9):1006–13.
- [5] ILAE. *EpilepsyDiagnosis.org*. <[www.epilepsydiagnosis.org](http://www.epilepsydiagnosis.org)>.
- [6] Fisher RS, Cross JH, D'Souza C, et al. Instruction manual for the ILAE 2017 operational classification of seizure types. *Epilepsia* 2017;58(4):531–42.
- [7] Cerminara C, Coniglio A, El-Malhany N, Casarelli L, Curatolo P. Two epileptic syndromes, one brain: childhood absence epilepsy and benign childhood epilepsy with centrotemporal spikes. *Seizure* 2012;21(1):70–4.
- [8] Hesdorffer DC, Benn EK, Cascino GD, Hauser WA. Is a first acute symptomatic seizure epilepsy? Mortality and risk for recurrent seizure. *Epilepsia* 2009;50(5):1102–8.
- [9] Fisher RS, Acevedo C, Arzimanoglou A, et al. ILAE official report: a practical clinical definition of epilepsy. *Epilepsia* 2014;55(4):475–82.
- [10] Hauser WA, Rich SS, Lee JR, Annegers JF, Anderson VE. Risk of recurrent seizures after two unprovoked seizures. *N Engl J Med* 1998;338(7):429–34.
- [11] Engel Jr. J. Report of the ILAE classification core group. *Epilepsia* 2006;47(9):1558–68.
- [12] GBD Epilepsy Collaborators, Global, regional, and national burden of epilepsy, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019;18(4):357–75.
- [13] Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: a 30-year longitudinal cohort study. *JAMA Neurol* 2018;75(3):279–86.
- [14] Scheffer IE, Berkovic S, Capovilla G, et al. ILAE classification of the epilepsies: position paper of the ILAE commission for classification and terminology. *Epilepsia* 2017;58(4):512–21.
- [15] Berg AT, Berkovic SF, Brodie MJ, et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE commission on classification and terminology, 2005–2009. *Epilepsia* 2010;51(4):676–85.
- [16] Vadlamudi L, Andermann E, Lombroso CT, et al. Epilepsy in twins: insights from unique historical data of William Lennox. *Neurology* 2004;62(7):1127–33.
- [17] Berkovic SF, Howell RA, Hay DA, Hopper JL. Epilepsies in twins: genetics of the major epilepsy syndromes. *Ann Neurol* 1998;43(4):435–45.
- [18] Kjeldsen MJ, Corey LA, Christensen K, Friis ML. Epileptic seizures and syndromes in twins: the importance of genetic factors. *Epilepsy Res* 2003;55(1–2):137–46.
- [19] Vadlamudi L, Milne RL, Lawrence K, et al. Genetics of epilepsy: the testimony of twins in the molecular era. *Neurology* 2014;83(12):1042–8.
- [20] Myers KA, Scheffer IE, Berkovic SF, Commission IG. Genetic literacy series: genetic epilepsy with febrile seizures plus. *Epileptic Disord* 2018;20(4):232–8.
- [21] Peljto AL, Barker-Cummings C, Vasoli VM, et al. Familial risk of epilepsy: a population-based study. *Brain* 2014;137(Pt 3):795–805.

- [22] Winawer MR, Rabinowitz D, Barker-Cummings C, et al. Evidence for distinct genetic influences on generalized and localization-related epilepsy. *Epilepsia* 2003;44(9):1176–82.
- [23] Epi4K Consortium. Phenotypic analysis of 303 multiplex families with common epilepsies. *Brain* 2017;140(8):2144–56.
- [24] Tobochnik S, Fahlstrom R, Shain C, Winawer MR, Investigators E. Familial aggregation of focal seizure semiology in the Epilepsy Phenome/Genome Project. *Neurology* 2017;89(1):22–8.
- [25] Winawer MR, Rabinowitz D, Pedley TA, Hauser WA, Ottman R. Genetic influences on myoclonic and absence seizures. *Neurology* 2003;61(11):1576–81.
- [26] Marini C, Scheffer IE, Crossland KM, et al. Genetic architecture of idiopathic generalized epilepsy: clinical genetic analysis of 55 multiplex families. *Epilepsia* 2004;45(5):467–78.
- [27] Winawer MR, Marini C, Grinton BE, et al. Familial clustering of seizure types within the idiopathic generalized epilepsies. *Neurology* 2005;65(4):523–8.
- [28] Kinirons P, Rabinowitz D, Gravel M, et al. Phenotypic concordance in 70 families with IGE-implications for genetic studies of epilepsy. *Epilepsy Res* 2008;82(1):21–8.
- [29] Koeleman BPC. What do genetic studies tell us about the heritable basis of common epilepsy? Polygenic or complex epilepsy? *Neurosci Lett* 2018;667:10–16.
- [30] Greenberg DA, Stewart WC. How should we be searching for genes for common epilepsy? A critique and a prescription. *Epilepsia* 2012;53(Suppl 4):72–80.
- [31] Epi Kc, Epilepsy Phenome/Genome P. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol* 2017;16(2):135–43.
- [32] May P, Girard S, Harrer M, et al. Rare coding variants in genes encoding GABAA receptors in genetic generalised epilepsies: an exome-based case-control study. *Lancet Neurol* 2018;17(8):699–708.
- [33] Pippucci T, Licchetta L, Baldassari S, et al. Contribution of ultrarare variants in mTOR pathway genes to sporadic focal epilepsies. *Ann Clin Transl Neurol* 2019;6(3):475–85.
- [34] Bisulli F, Tinuper P, Avoni P, et al. Idiopathic partial epilepsy with auditory features (IPEAF): a clinical and genetic study of 53 sporadic cases. *Brain* 2004;127(Pt 6):1343–52.
- [35] Liu H, Lu C, Li Z, et al. The identification of a novel mutation of nicotinic acetylcholine receptor gene CHRNA2 in a Chinese patient: Its possible implication in non-familial nocturnal frontal lobe epilepsy. *Epilepsy Res* 2011;95(1–2):94–9.
- [36] Baulac S, Gourfinkel-An I, Picard F, et al. A second locus for familial generalized epilepsy with febrile seizures plus maps to chromosome 2q21–q33. *Am J Hum Genet* 1999;65(4):1078–85.
- [37] Moulard B, Guipponi M, Chaigne D, Mouthon D, Buresi C, Malafosse A. Identification of a new locus for generalized epilepsy with febrile seizures plus (GEFS+) on chromosome 2q24–q33. *Am J Hum Genet* 1999;65(5):1396–400.
- [38] Escayg A, MacDonald BT, Meisler MH, et al. Mutations in SCN1A, encoding a neuronal sodium channel, in two families with GEFS + 2. *Nat Genet* 2000;24(4):343–5.
- [39] Ottman R, Risch N, Hauser WA, et al. Localization of a gene for partial epilepsy to chromosome 10q. *Nat Genet* 1995;10(1):56–60.
- [40] Winawer MR, Martinelli Boneschi F, Barker-Cummings C, et al. Four new families with autosomal dominant partial epilepsy with auditory features: clinical description and linkage to chromosome 10q24. *Epilepsia* 2002;43(1):60–7.
- [41] Kalachikov S, Evgrafov O, Ross B, et al. Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features. *Nat Genet* 2002;30(3):335–41.
- [42] Phillips HA, Scheffer IE, Berkovic SF, Hollway GE, Sutherland GR, Mulley JC. Localization of a gene for autosomal dominant nocturnal frontal lobe epilepsy to chromosome 20q 13.2. *Nat Genet* 1995;10(1):117–18.
- [43] Steinlein OK, Mulley JC, Propping P, et al. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* 1995;11(2):201–3.
- [44] Aridon P, Marini C, Di Resta C, et al. Increased sensitivity of the neuronal nicotinic receptor alpha 2 subunit causes familial epilepsy with nocturnal wandering and ictal fear. *Am J Hum Genet* 2006;79(2):342–50.
- [45] De Fusco M, Becchetti A, Patrignani A, et al. The nicotinic receptor beta 2 subunit is mutant in nocturnal frontal lobe epilepsy. *Nat Genet* 2000;26(3):275–6.
- [46] Gambardella A, Annesi G, De Fusco M, et al. A new locus for autosomal dominant nocturnal frontal lobe epilepsy maps to chromosome 1. *Neurology* 2000;55(10):1467–71.
- [47] Heron SE, Smith KR, Bahlo M, et al. Missense mutations in the sodium-gated potassium channel gene KCNT1 cause severe autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* 2012;44(11):1188–90.
- [48] Xiong L, Labuda M, Li DS, et al. Mapping of a gene determining familial partial epilepsy with variable foci to chromosome 22q11–q12. *Am J Hum Genet* 1999;65(6):1698–710.
- [49] Ishida S, Picard F, Rudolf G, et al. Mutations of DEPDC5 cause autosomal dominant focal epilepsies. *Nat Genet* 2013;45(5):552–5.
- [50] Lal D, Reinthaler EM, Schubert J, et al. DEPDC5 mutations in genetic focal epilepsies of childhood. *Ann Neurol* 2014;75(5):788–92.
- [51] Picard F, Makrythanasis P, Navarro V, et al. DEPDC5 mutations in families presenting as autosomal dominant nocturnal frontal lobe epilepsy. *Neurology* 2014;82(23):2101–6.
- [52] Baulac S, Ishida S, Marsan E, et al. Familial focal epilepsy with focal cortical dysplasia due to DEPDC5 mutations. *Ann Neurol* 2015;77(4):675–83.
- [53] Tsai MH, Chan CK, Chang YC, et al. DEPDC5 mutations in familial and sporadic focal epilepsy. *Clin Genet* 2017;92(4):397–404.
- [54] Strug LJ, Clarke T, Chiang T, et al. Centrottemporal sharp wave EEG trait in rolandic epilepsy maps to Elongator Protein Complex 4 (ELP4). *Eur J Hum Genet* 2009;17(9):1171–81.
- [55] Neubauer BA, Fiedler B, Himmelein B, et al. Centrottemporal spikes in families with rolandic epilepsy: linkage to chromosome 15q14. *Neurology* 1998;51(6):1608–12.
- [56] Lemke JR, Lal D, Reinthaler EM, et al. Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet* 2013;45(9):1067–72.

- [57] Chahine L, Abou-Khalil B, Siren A, et al. A new locus for familial temporal lobe epilepsy on chromosome 3q. *Epilepsy Res* 2013;106(3):338–44.
- [58] Fanciulli M, Di Bonaventura C, Egeo G, et al. Suggestive linkage of familial mesial temporal lobe epilepsy to chromosome 3q26. *Epilepsy Res* 2014;108(2):232–40.
- [59] Hedera P, Blair MA, Andermann E, et al. Familial mesial temporal lobe epilepsy maps to chromosome 4q13.2-q21.3. *Neurology* 2007;68(24):2107–12.
- [60] Azmanov DN, Zhelyazkova S, Radionova M, et al. Focal epilepsy of probable temporal lobe origin in a Gypsy family showing linkage to a novel locus on 7p21.3. *Epilepsy Res* 2011;96(1–2):101–8.
- [61] Claes L, Audenaert D, Deprez L, et al. Novel locus on chromosome 12q22-q23.3 responsible for familial temporal lobe epilepsy associated with febrile seizures. *J Med Genet* 2004;41(9):710–14.
- [62] Baulac S, Picard F, Herman A, et al. Evidence for digenic inheritance in a family with both febrile convulsions and temporal lobe epilepsy implicating chromosomes 18qter and 1q25-q31. *Ann Neurol* 2001;49(6):786–92.
- [63] Consortium E, Leu C, de Kovel CG, et al. Genome-wide linkage meta-analysis identifies susceptibility loci at 2q34 and 13q31.3 for genetic generalized epilepsies. *Epilepsia* 2012;53(2):308–18.
- [64] Durner M, Keddache MA, Tomasini L, et al. Genome scan of idiopathic generalized epilepsy: evidence for major susceptibility gene and modifying genes influencing the seizure type. *Ann Neurol* 2001;49(3):328–35.
- [65] Greenberg DA, Cayanis E, Strug L, et al. Malic enzyme 2 may underlie susceptibility to adolescent-onset idiopathic generalized epilepsy. *Am J Hum Genet* 2005;76(1):139–46.
- [66] Wang M, Greenberg DA, Stewart WCL. Replication, reanalysis, and gene expression: ME2 and genetic generalized epilepsy. *Epilepsia* 2019;60(3):539–46.
- [67] Greenberg DA, Delgado-Escueta AV, Wideltitz H, et al. Juvenile myoclonic epilepsy (JME) may be linked to the BF and HLA loci on human chromosome 6. *Am J Med Genet* 1988;31(1):185–92.
- [68] Suzuki T, Delgado-Escueta AV, Aguan K, et al. Mutations in EFHC1 cause juvenile myoclonic epilepsy. *Nat Genet* 2004;36(8):842–9.
- [69] Bailey JN, de Nijs L, Bai D, et al. Variant intestinal-cell kinase in juvenile myoclonic epilepsy. *N Engl J Med* 2018;378(11):1018–28.
- [70] Pathak S, Miller J, Morris EC, Stewart WCL, Greenberg DA. DNA methylation of the BRD2 promoter is associated with juvenile myoclonic epilepsy in Caucasians. *Epilepsia* 2018;59(5):1011–19.
- [71] Lenzen KP, Heils A, Lorenz S, Hempelmann A, Sander T. Association analysis of malic enzyme 2 gene polymorphisms with idiopathic generalized epilepsy. *Epilepsia* 2005;46(10):1637–41.
- [72] Subaran RL, Conte JM, Stewart WC, Greenberg DA. Pathogenic EFHC1 mutations are tolerated in healthy individuals dependent on reported ancestry. *Epilepsia* 2015;56(2):188–94.
- [73] Schulz H, Ruppert AK, Zara F, et al. No evidence for a BRD2 promoter hypermethylation in blood leukocytes of Europeans with juvenile myoclonic epilepsy. *Epilepsia* 2019;60(5):e31–6.
- [74] Lerche H, Berkovic SF, Lowenstein DH, Euro E-CC, Epi PGXC, Epi KCEGP. Intestinal-cell kinase and juvenile myoclonic epilepsy. *N Engl J Med* 2019;380(16):e24.
- [75] de Kovel CG, Pinto D, de Haan GJ, Kasteleijn-Nolst Trenite DG, Lindhout D, Koeleman BP. Association analysis of BRD2 (RING3) and epilepsy in a Dutch population. *Epilepsia* 2007;48(11):2191–2.
- [76] Cavalleri GL, Walley NM, Soranzo N, et al. A multicenter study of BRD2 as a risk factor for juvenile myoclonic epilepsy. *Epilepsia* 2007;48(4):706–12.
- [77] Kasperaviciute D, Catarino CB, Heinzen EL, et al. Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. *Brain* 2010;133(Pt 7):2136–47.
- [78] Kasperaviciute D, Catarino CB, Matarin M, et al. Epilepsy, hippocampal sclerosis and febrile seizures linked by common genetic variation around SCN1A. *Brain* 2013;136(Pt 10):3140–50.
- [79] International League Against Epilepsy Consortium on Complex Epilepsies, Electronic address: epilepsy-austin@unimelb.edu.au. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol* 2014;13(9):893–903.
- [80] International League Against Epilepsy Consortium on Complex E. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun* 2018;9(1):5269.
- [81] Speed D, O'Brien TJ, Palotie A, et al. Describing the genetic architecture of epilepsy through heritability analysis. *Brain* 2014;137(Pt 10):2680–9.
- [82] McTague A, Howell KB, Cross JH, Kurian MA, Scheffer IE. The genetic landscape of the epileptic encephalopathies of infancy and childhood. *Lancet Neurol* 2016;15(3):304–16.
- [83] Berg AT, Coryell J, Saneto RP, et al. Early-life epilepsies and the emerging role of genetic testing. *JAMA Pediatr* 2017;171(9):863–71.
- [84] Epi4K Consortium, Epilepsy Phenome/Genome Project, Allen AS, et al. De novo mutations in epileptic encephalopathies. *Nature* 2013;501(7466):217–21.
- [85] EuroEPINOMICS-RES Consortium, Epilepsy Phenome/Genome Project, Epi4K Consortium. De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. *Am J Hum Genet* 2014;95(4):360–70.
- [86] Epi4K Consortium. De novo mutations in SLC1A2 and CACNA1A are important causes of epileptic encephalopathies. *Am J Hum Genet* 2016;99(2):287–98.
- [87] EuroEPINOMICS-RES Consortium, Epilepsy Phenome/Genome Project, Epi4K Consortium. De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. *Am J Hum Genet* 2017;100(1):179.
- [88] Heyne HO, Singh T, Stamberger H, et al. De novo variants in neurodevelopmental disorders with epilepsy. *Nat Genet* 2018;50(7):1048–53.
- [89] Takata A, Nakashima M, Saito H, et al. Comprehensive analysis of coding variants highlights genetic complexity in developmental and epileptic encephalopathy. *Nat Commun* 2019;10(1):2506.
- [90] Hamdan FF, Myers CT, Cossette P, et al. High rate of recurrent de novo mutations in developmental and epileptic encephalopathies. *Am J Hum Genet* 2017;101(5):664–85.



- [91] Dibbens LM, Tarpey PS, Hynes K, et al. X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. *Nat Genet* 2008;40(6):776–81.
- [92] Calhoun JD, Carvill GL. Unravelling the genetic architecture of autosomal recessive epilepsy in the genomic era. *J Neurogenet* 2018;32(4):295–312.
- [93] Myers CT, Hollingsworth G, Muir AM, et al. Parental mosaicism in “de novo” epileptic encephalopathies. *N Engl J Med* 2018;378(17):1646–8.
- [94] Mercimek-Mahmutoglu S, Patel J, Cordeiro D, et al. Diagnostic yield of genetic testing in epileptic encephalopathy in childhood. *Epilepsia* 2015;56(5):707–16.
- [95] Helbig KL, Farwell Hagman KD, Shinde DN, et al. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. *Genet Med* 2016;18(9):898–905.
- [96] Klassen T, Davis C, Goldman A, et al. Exome sequencing of ion channel genes reveals complex profiles confounding personal risk assessment in epilepsy. *Cell* 2011;145(7):1036–48.
- [97] Heinzen EL, Depondt C, Cavalleri GL, et al. Exome sequencing followed by large-scale genotyping fails to identify single rare variants of large effect in idiopathic generalized epilepsy. *Am J Hum Genet* 2012;91(2):293–302.
- [98] Bobbili DR, Lal D, May P, et al. Exome-wide analysis of mutational burden in patients with typical and atypical rolandic epilepsy. *Eur J Hum Genet* 2018;26(2):258–64.
- [99] Lalioti MD, Scott HS, Genton P, et al. A PCR amplification method reveals instability of the dodecamer repeat in progressive myoclonus epilepsy (EPM1) and no correlation between the size of the repeat and age at onset. *Am J Hum Genet* 1998;62(4):842–7.
- [100] Parrini E, Ramazzotti A, Dobyns WB, et al. Periventricular heterotopia: phenotypic heterogeneity and correlation with Filamin A mutations. *Brain* 2006;129(Pt 7):1892–906.
- [101] Parrini E, Mei D, Wright M, Dorn T, Guerrini R. Mosaic mutations of the FLN1 gene cause a mild phenotype in patients with periventricular heterotopia. *Neurogenetics* 2004;5(3):191–6.
- [102] Zhang X, Chen MH, Wu X, et al. Cell-type-specific alternative splicing governs cell fate in the developing cerebral cortex. *Cell* 2016;166(5):1147–1162.e15.
- [103] Carvill GL, Engel KL, Ramamurthy A, et al. Aberrant inclusion of a poison exon causes Dravet syndrome and related SCN1A-associated genetic epilepsies. *Am J Hum Genet* 2018;103(6):1022–9.
- [104] Ishiura H, Doi K, Mitsui J, et al. Expansions of intronic TTTCa and TTTTA repeats in benign adult familial myoclonic epilepsy. *Nat Genet* 2018;50(4):581–90.
- [105] Verrotti A, Sertorio F, Matricardi S, Ferrara P, Striano P. Electroclinical features of epilepsy in patients with InvDup(15). *Seizure* 2017;47:87–91.
- [106] Matricardi S, Darra F, Spalice A, et al. Electroclinical findings and long-term outcomes in epileptic patients with inv dup (15). *Acta Neurol Scand* 2018;137(6):575–81.
- [107] Daber RD, Conlin LK, Leonard LD, et al. Ring chromosome 20. *Eur J Med Genet* 2012;55(5):381–7.
- [108] Unterberger I, Dobesberger J, Schober H, et al. A further case of familial ring chromosome 20 mosaicism—molecular characterization of the ring and review of the literature. *Eur J Med Genet* 2018; Available from: <https://doi.org/10.1016/j.ejmg.2018.10.016>.
- [109] Kendall KM, Rees E, Escott-Price V, et al. Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK biobank subjects. *Biol Psychiatry* 2017;82(2):103–10.
- [110] Balciuniene J, DeChene ET, Akgumus G, et al. Use of a dynamic genetic testing approach for childhood-onset epilepsy. *JAMA Netw Open* 2019;2(4):e192129.
- [111] Leu C, Coppola A, Sisodiya SM. Progress from genome-wide association studies and copy number variant studies in epilepsy. *Curr Opin Neurol* 2016;29(2):158–67.
- [112] Mullen SA, Berkovic SF, Commission IG. Genetic generalized epilepsies. *Epilepsia* 2018;59(6):1148–53.
- [113] Lauschke VM, Milani L, Ingelman-Sundberg M. Pharmacogenomic biomarkers for improved drug therapy—recent progress and future developments. *AAPS J* 2017;20(1):4.
- [114] de Kovel CG, Trucks H, Helbig I, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 2010;133(Pt 1):23–32.
- [115] Helbig I, Mefford HC, Sharp AJ, et al. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet* 2009;41(2):160–2.
- [116] Mefford HC, Muhle H, Ostertag P, et al. Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. *PLoS Genet* 2010;6(5):e1000962.
- [117] Perez-Palma E, Helbig I, Klein KM, et al. Heterogeneous contribution of microdeletions in the development of common generalised and focal epilepsies. *J Med Genet* 2017;54(9):598–606.
- [118] Moller RS, Weber YG, Klitten LL, et al. Exon-disrupting deletions of NRXN1 in idiopathic generalized epilepsy. *Epilepsia* 2013;54(2):256–64.
- [119] Lal D, Ruppert AK, Trucks H, et al. Burden analysis of rare microdeletions suggests a strong impact of neurodevelopmental genes in genetic generalised epilepsies. *PLoS Genet* 2015;11(5):e1005226.
- [120] Lal D, Trucks H, Moller RS, et al. Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. *Epilepsia* 2013;54(2):265–71.
- [121] Addis L, Rosch RE, Valentin A, et al. Analysis of rare copy number variation in absence epilepsies. *Neurol Genet* 2016;2(2):e56.
- [122] Jabbari K, Bobbili DR, Lal D, et al. Rare gene deletions in genetic generalized and rolandic epilepsies. *PLoS One* 2018;13(8):e0202022.
- [123] Heinzen EL, Radtke RA, Urban TJ, et al. Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. *Am J Hum Genet* 2010;86(5):707–18.
- [124] Striano P, Coppola A, Paravidino R, et al. Clinical significance of rare copy number variations in epilepsy: a case-control survey using microarray-based comparative genomic hybridization. *Arch Neurol* 2012;69(3):322–30.

- [125] Mullen SA, Carvill GL, Bellows S, et al. Copy number variants are frequent in genetic generalized epilepsy with intellectual disability. *Neurology* 2013;81(17):1507–14.
- [126] Mefford HC, Yendle SC, Hsu C, et al. Rare copy number variants are an important cause of epileptic encephalopathies. *Ann Neurol* 2011;70(6):974–85.
- [127] Epilepsy Phenome/Genome Project Epi KC. Copy number variant analysis from exome data in 349 patients with epileptic encephalopathy. *Ann Neurol* 2015;78(2):323–8.
- [128] Helbig I, Swinkels ME, Aten E, et al. Structural genomic variation in childhood epilepsies with complex phenotypes. *Eur J Hum Genet* 2014;22(7):896–901.
- [129] Monlong J, Girard SL, Meloche C, et al. Global characterization of copy number variants in epilepsy patients from whole genome sequencing. *PLoS Genet* 2018;14(4):e1007285.
- [130] Steele HE, Chinnery PF. Mitochondrial causes of epilepsy: evaluation, diagnosis, and treatment. *Semin Neurol* 2015;35(3):300–9.
- [131] Ortigoza-Escobar JD, Molero-Luis M, Arias A, et al. Treatment of genetic defects of thiamine transport and metabolism. *Expert Rev Neurother* 2016;16(7):755–63.
- [132] Malek N, Stewart W, Greene J. The progressive myoclonic epilepsies. *Pract Neurol* 2015;15(3):164–71.
- [133] Mullan KA, Anderson A, Illing PT, Kwan P, Purcell AW, Mifsud NA. HLA-associated antiepileptic drug-induced cutaneous adverse reactions. *HLA* 2019;93(6):417–35.
- [134] Gatanaga H, Honda H, Oka S. Pharmacogenetic information derived from analysis of HLA alleles. *Pharmacogenomics* 2008;9(2):207–14.
- [135] Ferrell PB, McLeod HL. Carbamazepine, HLA-B\*1502 and risk of Stevens-Johnson syndrome and toxic epidermal necrolysis: US FDA recommendations. *Pharmacogenomics* 2008;9(10):1543–6.
- [136] Jaruthamsophon K, Tipmanee V, Sangiemchoey A, Sukasem C, Limprasert P. HLA-B\*15:21 and carbamazepine-induced Stevens-Johnson syndrome: pooled-data and in silico analysis. *Sci Rep* 2017;7:45553.
- [137] Silvado CE, Terra VC, Twardowsky CA. CYP2C9 polymorphisms in epilepsy: influence on phenytoin treatment. *Pharmacogenomics Pers Med* 2018;11:51–8.
- [138] Bajpai M, Roskos LK, Shen DD, Levy RH. Roles of cytochrome P4502C9 and cytochrome P4502C19 in the stereoselective metabolism of phenytoin to its major metabolite. *Drug Metab Dispos* 1996;24(12):1401–3.
- [139] Kidd RS, Curry TB, Gallagher S, Edeki T, Blaisdell J, Goldstein JA. Identification of a null allele of CYP2C9 in an African-American exhibiting toxicity to phenytoin. *Pharmacogenetics* 2001;11(9):803–8.
- [140] Depondt C, Godard P, Espel RS, Da Cruz AL, Lienard P, Pandolfo M. A candidate gene study of antiepileptic drug tolerability and efficacy identifies an association of CYP2C9 variants with phenytoin toxicity. *Eur J Neurol* 2011;18(9):1159–64.
- [141] McCormack M, Gui H, Ingason A, et al. Genetic variation in CFH predicts phenytoin-induced maculopapular exanthema in European-descent patients. *Neurology* 2018;90(4):e332–41.
- [142] Verbelen M, Weale ME, Lewis CM. Cost-effectiveness of pharmacogenetic-guided treatment: are we there yet? *Pharmacogenomics J* 2017;17(5):395–402.
- [143] Marson AG, Al-Kharusi AM, Alwaidh M, et al. The SANAD study of effectiveness of carbamazepine, gabapentin, lamotrigine, oxcarbazepine, or topiramate for treatment of partial epilepsy: an unblinded randomised controlled trial. *Lancet* 2007;369(9566):1000–15.
- [144] Kwan P, Poon WS, Ng HK, et al. Multidrug resistance in epilepsy and polymorphisms in the voltage-gated sodium channel genes SCN1A, SCN2A, and SCN3A: correlation among phenotype, genotype, and mRNA expression. *Pharmacogenet Genomics* 2008;18(11):989–98.
- [145] Haerian BS, Baum L, Kwan P, Tan HJ, Raymond AA, Mohamed Z. SCN1A, SCN2A and SCN3A gene polymorphisms and responsiveness to antiepileptic drugs: a multicenter cohort study and meta-analysis. *Pharmacogenomics* 2013;14(10):1153–66.
- [146] Bao Y, Liu X, Xiao Z. Association between two SCN1A polymorphisms and resistance to sodium channel blocking AEDs: a meta-analysis. *Neurol Sci* 2018;39(6):1065–72.
- [147] Wolff M, Johannesen KM, Hedrich UBS, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. *Brain* 2017;140(5):1316–36.
- [148] Berghuis B, Stapleton C, Sonnsma ACM, et al. A genome-wide association study of sodium levels and drug metabolism in an epilepsy cohort treated with carbamazepine and oxcarbazepine. *Epilepsia Open* 2019;4(1):102–9.
- [149] Howell KB, Eggers S, Dalziel K, et al. A population-based cost-effectiveness study of early genetic testing in severe epilepsies of infancy. *Epilepsia* 2018;59(6):1177–87.
- [150] Lindy AS, Stosser MB, Butler E, et al. Diagnostic outcomes for genetic testing of 70 genes in 8565 patients with epilepsy and neurodevelopmental disorders. *Epilepsia* 2018;59(5):1062–71.
- [151] Wang J, Lin ZJ, Liu L, et al. Epilepsy-associated genes. *Seizure* 2017;44:11–20.
- [152] Nashef L, Singh R, Moran N, Murphy E. Investigating adults with early-onset epilepsy and intellectual or physical disability. *Pract Neurol* 2019;19(2):115–30.
- [153] Perucca P, Perucca E. Identifying mutations in epilepsy genes: Impact on treatment selection. *Epilepsy Res* 2019;152:18–30.
- [154] Epilepsy Genetics Initiative. The Epilepsy Genetics Initiative: systematic reanalysis of diagnostic exomes increases yield. *Epilepsia* 2019;60:797–806.
- [155] Pal D, Helbig I. Commentary: pathogenic EFHC1 mutations are tolerated in healthy individuals dependent on reported ancestry. *Epilepsia* 2015;56(2):195–6.
- [156] Leen WG, Klepper J, Verbeek MM, et al. Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. *Brain* 2010;133(Pt 3):655–70.
- [157] Liu YC, Lee JW, Bellows ST, et al. Evaluation of non-coding variation in GLUT1 deficiency. *Dev Med Child Neurol* 2016;58(12):1295–302.
- [158] Zaman SM, Mullen SA, Petrovski S, et al. Development of a rapid functional assay that predicts GLUT1 disease severity. *Neurol Genet* 2018;4(6):e297.



# The human leukocyte antigen system in human disease and transplantation medicine

Gaurav Sharma<sup>1</sup>, Ajay Kumar Baranwal<sup>2</sup> and Narinder Kumar Mehra<sup>3</sup>

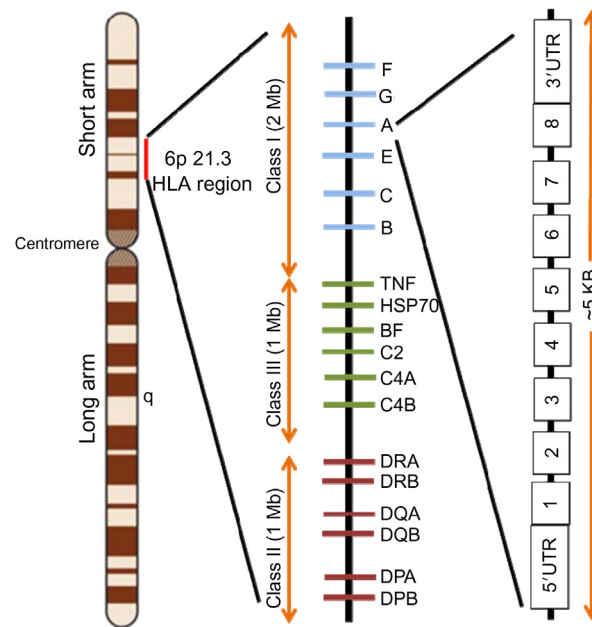
<sup>1</sup>Department of Translational and Regenerative Medicine, Postgraduate Institute of Medical Education and Research, Chandigarh, India <sup>2</sup>Command Hospital - Armed Forces Medical College, Pune, India <sup>3</sup>National Chair and Former Dean (Research), All India Institute of Medical Sciences, New Delhi, India

## 17.1 Introduction

Major histocompatibility complex (MHC) region is characterized by multiple genetic loci that play a major role in the development of both cell-mediated and humoral immune responses and also determines whether a transplanted tissue will be accepted as self or rejected as foreign. The discovery of MHC is credited to the initial transplantation studies in mice, which suggested that an immune response to cell-surface antigens result in rejection of foreign tissue. Particularly, in the mid-1930s, Peter Gorer from Lister Institute, London, used inbred mice to identify these blood group antigens and defined four groups of genes (I–IV). Later in 1940s he collaborated with George Snell from Jackson Laboratories, the United States and established that antigens encoded by the genes in group II resulted in the rejection of transplanted tumors and foreign tissue. These genes were designated as “histocompatibility genes” (H-2 genes) by Snell, in reference to Gorer’s blood group II antigens. Though Gorer died before the importance of his work was recognized, Snell was awarded the Nobel Prize for this work in 1980, which he shared with Jean Dausset (France) for describing the first MHC antigen in humans [membrane attack complex (MAC) or human leukocyte antigen (HLA)-A2] and Baruj Benacerraf (United States) for proposing role of immune response genes (Ir) in MHC-mediated immune responses. This work was followed by the discovery of 4a and 4b (HLA-Bw4 and Bw6) by Jon van Rood et al. in The Netherlands. Over the years, several other prominent investigators immensely contributed to this field including Rolf Zinkernagel and Peter Doherty who shared the Nobel Prize in 1996 for their discovery of MHC restriction and associated mechanisms.

In humans the collection of MHC genes arrayed within this region spans  $4 \times 10^6$  nucleotides on the short arm of chromosomes 6 at position 6p21.3 (Fig. 17.1). It is the most polymorphic region of the human genome that encodes for HLA system. As of February 2019, a total of ~21,499 HLA alleles have been named, and several of their protein variants are expanding with the time (<http://www.ebi.ac.uk/imgt/hla/stats.html> accessed 24.02.19). Indeed, the extreme polymorphism of this region across various ethnic groups suggests its evolutionary significance and the selection pressure at this region. This polymorphism could be attributed to several factors that include (1) high mutation rate (nonsynonymous/synonymous), (2) overdominant selection that can be described as heterozygote advantage, that is, higher fitness of heterozygous genotype over homozygous, (3) frequency-based selection, an evolutionary process through which the fitness of a phenotype depends on its frequency relative to other phenotypes in a population. The fitness of a phenotype can increase (positive) or decrease (negative) as it becomes more common, and (4) gene conversion, that is, conversion of one allele to other during meiosis, which is attributed to homologous recombination between heterozygotic sites resulting in a mismatch in base pairing.

Besides its overly influence in controlling the long-term graft survival following organ and hematopoietic stem-cell transplantation, research in the field of MHC has progressed to reveal new mechanistic insights into the pathogenesis of several diseases, in particular those with infectious and autoimmune etiology. Based on its biological



**FIGURE 17.1** Human MHC with chromosomal location and gene map showing multiple genes on the short arm of the chromosome 6 (6p21.3). The two-way arrow shows the genetic distance covered by respective regions on the chromosome. *MHC*, Major histocompatibility complex.

relevance, the HLA system can be considered a “mini genome model.” More appropriately, it can be designated as “self-surveillance complex” as its relevance in health and disease is beyond defining the histocompatibility alone. Some of the inherent features that distinguish the human MHC from rest of the genome are summarized in [Box 17.1](#).

## 17.2 Human leukocyte antigen system

Human MHC contains a set of highly polymorphic genes, which includes HLA class I (HLA-A, B, C) and class II (HLA-DR, DQ, DP) regions, separated by MHC class III region that contains genes, which are directly or indirectly involved in the immune response, for example, complement genes (C4, C2, and Bf), cytokine genes (TNF- $\alpha$ , LTA, LTB), and several other immune-associated genes that include MICA, MICB, and heat shock protein genes ([Fig. 17.1](#)).

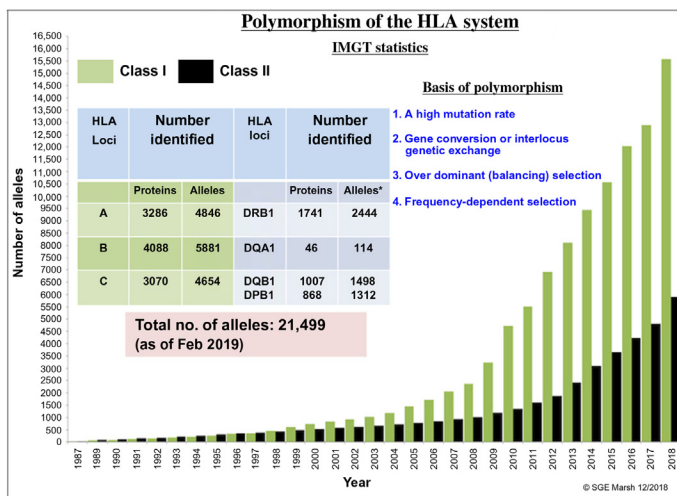
Data from physical mapping, sequencing, and cloning has revealed that a large number of genes are present in three regions: class I, class II, and class III. The most centromeric segment is the “class II region” containing the HLA-DR, DQ, and DP loci and spans around 1 Mb. At the telomeric end lies the “class I region” that is spread over a region of 2 Mb and contains the classical HLA-A, B, and C and related loci. The class I and class II regions show considerable homologies to one another and also exert the major transplantation influence through classical transplantation loci. Matching donor and recipient pairs for organ transplantation for alleles in these regions is therefore critical to long-term graft survival. All of these genes are situated so close together that they are usually inherited en bloc as a haplotype from either parent. The HLA system thus can be considered a single genetic unit.

**HLA class I molecules:** HLA class I molecules are heterodimeric glycoproteins composed of two polypeptide chains; alpha or heavy chain encoded by the MHC region and non-MHC-encoded light chain ( $\beta$ 2-microglobulin). The MHC-encoded polypeptide is about 350 amino acids long with a total molecular weight of about 45 kDa divided into three functional regions; external, transmembrane, and intracytoplasmic. The extracellular portion of the heavy chain is folded into three globular domains  $\alpha$ 1,  $\alpha$ 2, and  $\alpha$ 3, each of which contains stretches of about 90 amino acids encoded by separate exons. While  $\alpha$ 1 and  $\alpha$ 2 domains take part in peptide binding, the  $\alpha$ 3 is essentially conserved. Currently 13,680 HLA class I alleles have been described, which includes 4846 at locus A, 5881 at locus B, 4654 at locus C, 27 at locus E, 38 at locus F, and 61 at locus G (<http://hla.alleles.org/nomenclature/stats.html> accessed 24.02.19).

## BOX 17.1

## Inherent features that distinguish the human MHC from rest of the genome

1. It represents the most genes dense region of the human genome encoding  $\geq 252$  expressed loci, including key immune response genes.
2. Most highly polymorphic genetic region at the nucleotide level reflecting highest trait-associated variant density. As of February 2019, a total of 21,499 HLA alleles have been named including 15,586 class I and 5913 class II alleles. For details, please refer to Fig. 17.2.
3. The linkage disequilibrium among its loci is very high due to lower recombination rates than the rest of the genome, as represented by conserved extended haplotypes.
4. It represents strongest *trans* e-QTL (expression quantitative trait loci) and meQTL (methylation QTL) in the genome. A QTL is a genomic locus that correlates with variation of a quantitative trait in the phenotype of an organism.
5. Alternative splicing is represented at a higher than average rate, attributed to sequence diversity.
6. Human MHC has highest gene expression levels across the genome and highest heritability of gene expression levels.
7. Ability to present antigenic peptides promiscuously to generate immune responses.



**FIGURE 17.2** The graph indicates the total number of HLA class I and class II alleles identified and held in the IPD-IMGT/HLA database since 1987 till December 2018. HLA, Human leukocyte antigen.

**HLA class II molecules:** The MHC class II molecule is a dimer of dimers constituted by two noncovalently associated glycoprotein chains called  $\alpha$  and  $\beta$ . Each chain is made up of two domains designated as  $\alpha 1$ ,  $\alpha 2$  and  $\beta 1$ ,  $\beta 2$ , respectively. The  $\alpha 1$  and  $\beta 1$  domains of the class II molecules form the peptide-binding groove and are structurally homologous to the  $\alpha 1$  and  $\alpha 2$  domains of the class I heavy chain. As in the class I molecule, maximum polymorphism lies in the groove formed by these domains. The membrane proximal  $\alpha 2$  and  $\beta 2$  correspond to the  $\alpha 3$  and  $\beta 2$  microglobulin of class I and are relatively conserved. Currently, 5913 HLA class II alleles have been described, which includes 2444 at locus DRB1, 1498 at locus DQB1, and 1312 at locus DPB1 (<http://hla.alleles.org/nomenclature/stats.html> accessed 24.02.19).

**Immunobiological functions:** HLA is codominantly expressed in every individual, and therefore each individual expresses two protein products of each locus at the cell surface. Because of the close linkage between HLA genes (recombination frequency of  $\sim 0.5\%$ ), the genes at each locus on a single chromosome are usually inherited as a haplotype, and each individual has two haplotypes. There is thus 25% chance that two siblings will share both haplotypes and will be fully HLA compatible, 50% chance that they will share one haplotype, and 25% chance that they will share no haplotype and thus will be completely HLA incompatible.

The main biological function of MHC molecule is to present antigens (peptides) to immune-competent cells. Within the antigen-presenting cells (APCs), antigens are degraded into small peptides, which in turn are linked to MHC molecules in order to express the peptide/MHC complex on its cell surface. This is essential for the efficient immune surveillance and recognition of alloantigen by lymphocytes. This entire process of antigen

presentation and processing is accomplished by sophisticated cellular machinery using a blend of cellular house-keeping proteins and dedicated transporters, chaperones, and peptidases. The preferential nature of peptide binding to class I or class II MHC molecules depends upon how the antigen enters into the cell. Antigens that are produced within cytoplasm (endogenous) of the APC tend to associate with class I MHC molecule for presentation to CD8 + cytotoxic T lymphocytes. On the other hand, exogenous antigens (those which are produced outside the cell), for example, extracellular bacterial products, viruses, parasites, or mismatched transplant antigens that are taken up by the APCs by phagocytosis, endocytosis, or exocytosis and tend to associate with the MHC class II molecules and are recognized by CD4 + helper T cells. Cross presentation is a phenomenon where exogenous antigens can be presented on MHC class I molecules and stimulate CD8 + T-cell immunity. Dendritic cells are the principal cells endowed with the capacity to cross-present antigens; however, it is not the only cell to do so. Macrophages and B cells can also cross-present. Antigens that can be cross-presented include cellular antigens, soluble proteins, immune complexes, intracellular bacteria, parasites, viral antigens, as well as normal tissue antigens. Cross-presented antigens are internalized through pinocytosis, receptor-mediated endocytosis and phagocytosis. Phagocytosis is particularly relevant in organ transplantation where recipients immunologically respond to donor antigens.

This chapter is an attempt to highlight specific aspects of the clinical and molecular importance of HLA as an important biomarker for human diseases and its relevance in organ transplantation. These aspects are covered in details in the following sections.

### 17.3 Human leukocyte antigen and disease

In recent years, candidate gene-based approaches have marked the advent of an era of rapidly evolving domain of “HLA and disease associations.” Since the first report in 1973 by Brewerton et al. in the United Kingdom, revealing a strong association linking ankylosing spondylitis with the presence of HLA-B27, candidate gene-based studies conducted in various populations have identified a long list of diseases associated with particular HLA alleles (<http://www.hladiseaseassociations.com/> accessed 15.01.19). Some of the most prominent associations include (1) ankylosing spondylitis and related spondyloarthropathies with specific HLA-B27 subtypes (HLA-B\*27:02, B\*27:04, B\*27:05), where >90% of patients carry these alleles; (2) narcolepsy, a disease with very high occurrence of DQB1\*06:02 allele in nearly 90%–100% of patients; (3) >90% of celiac disease patients harbor HLA-DQ2 molecule (HLA-DQA1\*05:01-DQB1\*02:01) while the remaining patients carry DQ8 molecule (HLA-DQA1\*03:01-DQB1\*03:02); (4) rheumatoid arthritis (RA), in which >90% Caucasian seropositive patients carry one or two HLA-DRB1 alleles coding a five amino acid coding sequence motif in the DRβ chain called “shared epitope” (e.g., DRB1\*04:01, DRB1\*04:04, DRB1\*04:05, and others); (5) type 1 diabetes (T1D) in which >90% patients carry either HLA-DRB1\*03/DQB1\*02:01 or DRB1\*04/DQB1\*03:02 or both haplotypes. These approaches have revealed definitive population-based differences and strongest association of HLA class II region genes with autoimmune and infectious diseases (Table 17.1).

Despite several attempts, exact mechanism(s) of most of the observed HLA and disease associations are largely unknown due to several reasons, which include (1) the extended linkage disequilibrium within the HLA region whereby the observed associations could be due to a combination of different HLA molecules expressed by different loci rather than a single HLA allele; (2) multifactorial and/or polygenic nature of the diseases, involving other genetic variants and environmental factors; (3) since the disease-specific autoantigens are largely unknown, this restricts the structural analysis of HLA–peptide interactions in autoimmune diseases as a causative link; (4) the observed appreciable heterogeneity in clinical presentation of most diseases, for example, age at onset, linkage to other diseases, variable response to therapy, clinical heterogeneity, and others.

Currently, several hypotheses have been proposed to explain the mechanistic basis of these associations. These can broadly be discussed under two categories: (1) “mistaken identity” according to which the observed HLA association could actually be due to another unknown locus linked to the associated allele. The specific HLA gene may have physical proximity to the region that might carry a gene conferring susceptibility or resistance to a particular disease. This hypothesis indicates the relevance of possible linkage disequilibrium, which is very high in this region. (2) The associated HLA alleles could act directly as disease vulnerability factors by means of antigenic cross-reactivity or mimicry of infectious agents with the particular HLA molecule (molecular mimicry) or possible immune response to self-antigens due to aberrant T-cell reactivity, cross-reactivity with foreign antigens or immune response to “altered self” antigens. Further, genetic susceptibility could largely or partly be due to the involvement of hitherto unknown non-HLA genes. Although various classical genetic studies in humans and experimental models have clearly documented the primary influence of MHC genes, data about the possible

**TABLE 17.1** List summarizing some of the prominent human leukocyte antigen (HLA) association with diseases across various populations.

S. no.	HLA allele/s	Disease
1	B*27:01, B*27:02, B*27:04, B*27:05	Ankylosing spondylitis
2	DRB1*03:01-DQB1*02 and DRB1*04-DQB1*03:02	Addison's disease
3	B5/B*51, particularly B*51:01	Behcet's disease
4	DQA1*05:01–DQB1*02:01, DQA1*02:01–DQB1*02:02, DQA1*03:01–DQB1*03:02	Celiac disease
5	DRB1*07, DRB1*01:03	Crohn's disease
6	DRB1*03:01-DQA1*05:01-DQB1*02:01 and DRB1*04:01-DQA1*03:01-DQB1*03:02	Graves' disease
7	HLA-B*27, B*57 (slow), B*35(Px) (fast)	HIV disease progression
8	DRB1*03:01, DQA1*05:01, DQB1*02:01 or DRB1*04:01, DQA1*03:01, DQB1*03:02	Type 1 diabetes
9	HLA-DRB1*15:01 on DRB1*15:01-DQA1*01:02-DQB1*06:02	Multiple sclerosis
10	DQB1*06:02, DRB1*15:01-DQA1*01:02-DQB1*06:02	Narcolepsy
11	C*06	Psoriasis vulgaris
12	DRB1*01:01, DRB1*01:02, DRB1*04:01, DRB1*04:04, DRB1*04:05, DRB1*04:08, DRB1*10:01, DRB1*13:03, DRB1*14:02 and DRB1*14:06	Rheumatoid arthritis
13	DRB1*11:04 and DQB1*03:01, DRB1*15:02 and DQB1*06:01, DQB1*05:01, DRB1*13:02 and DQB1*06:04	Scleroderma
14	DRB1*03–DQB1*02–DQA1*05:01; DRB1*15–DQB1*06–DQA1*01:02	Sjogren syndrome
15	DRB1*03:01-DQB1*02:01, DRB1*15:01-DQB1*06:02	Systemic lupus erythematosus

involvement of non-HLA genes in conferring susceptibility or resistance to various diseases is insufficient and beyond the scope of this chapter.

### 17.3.1 Human leukocyte antigen and drug-induced hypersensitivities

A few well-characterized examples of a strong association of specific HLA alleles with the development of drug-induced hypersensitivity reactions have been reported in the literature. These data have encouraged the development of pharmacological tests to screen patients at an obvious high risk of HLA-linked drug hypersensitivity reactions. Some of the examples include (1) HLA-B\*57:01-associated abacavir induced drug hypersensitive reaction (DHR) that has been observed in patients of HIV/AIDS [1]; (2) HLA-B\*58:01 with allopurinol induced hypersensitivity [2]; (3) HLA-B\*13:01-associated dapsone hypersensitivity [3]; and (4) HLA-B\*15:02-associated carbamazepine induced Stevens–Johnson syndrome/toxic epidermal necrolysis [4].

These HLA-linked DHRs can be attributed to three major mechanisms or models [5], each of which need to be explored further.

1. **Altered repertoire mechanism:** According to this hypothesis, the drug in question binds to antigen-binding cleft of the MHC molecule and alters the repertoire of HLA-bound peptides leading to presentation of self-peptides. For example, administration of abacavir has been shown to alter the repertoire of peptides bound to HLA-B\*57:01. X-ray crystallographic structures have revealed that the drug interacts with peptide and residues of antigen-binding cleft noncovalently.
2. **Hapten (prohapten) mechanism:** According to this hypothesis, the drugs or their metabolites form adducts with self-proteins resulting in immune recognition of hapten–self-peptide complexes, acting as de novo antigens. For example, penicillin can react chemically with proteins in the body and form hapten carrier complex that can cause anaphylaxis.
3. **Pharmacologic interaction (pi) mechanism:** As per this, drug can reversely and noncovalently bind to HLA or T-cell receptor molecules without binding to the antigen-binding cleft of the MHC and influence the immune response. Although no definitive HLA associations have been uncovered supporting this hypothesis, the same need to be explored further.



### 17.3.2 Epistatic interaction of major histocompatibility complex genes

Besides the direct involvement of HLA allelic variants, compelling evidence has been put forward to suggest the involvement of single nucleotide polymorphism (SNP) located outside the MHC region and having epistatic interaction with variants within MHC region, thus influencing disease outcome. These combinatorial associations add another layer of complexity to explain genetic involvement in disease outcome. Some of the examples include (1) the epistatic interaction among susceptibility loci in MHC with genetic polymorphisms in CTLA4 and IRF5 genes in systemic lupus erythematosus (SLE) [6], (2) HLA-B\*51 with endoplasmic reticulum aminopeptidase (ERAP1) in Behcet disease [7], (3) HLA and ERAP1 in ankylosing spondylitis [8], (4) HLA-C and ERAP1 in inflammatory bowel disease [9], (5) HLA and killer immunoglobulin-like receptors (KIR) interactions in HIV disease progression [10].

With the turn of the century the focus has shifted to genome-wide association studies (GWAS), involving the whole genome analysis. Indeed, the human MHC region continues to be the most strongly disease-associated genomic region with a number of diseases. Representative HLA and disease associations and corresponding GWAS are summarized by Kennedy et al. [11], where an attempt has been made to compile the data from National Human Genome Research Institute GWAS catalog ( $N = 18,682$ ).

Nevertheless the molecular mechanisms underlying HLA and disease associations remain largely unknown with limited progress in this direction. Specific examples include autoimmune diseases such as celiac disease, T1D [12], RA [13], and DHRs [5]. In celiac disease, ingested gluten is digested into amino acids and to proline-rich peptides, which are resistant to proteolytic cleavage. These undigested peptides are transported through the intestinal epithelium, deamidated by tissue transglutaminase in the mucosa and presented by HLA-DQ2 or DQ8 molecules on APCs, which are recognized by specific CD4T cells in the lamina propria, resulting in immune activation cascade. Similarly, in T1D, it has been shown that HLA-DR4 restricted CD4T cells specific to insulin A chain amino acid residues covered a major proportion of T cells in pancreatic lymph nodes of T1D patients. However, considering the large number of HLA disease associations, overall the literature revealing mechanistic insights has been limited. The following section includes some of the recent advances in this field.

## 17.4 Human leukocyte antigen expression: an explanation for disease development

In addition to the polymorphic genetic diversity of the HLA system and its association with various diseases at the candidate gene and genome-wide levels, another layer of complexity includes polymorphic allele-specific expression. Although research in this promising field is still in nascent stage, the expression variables in different disease could help in defining the mechanism underlying HLA-linked disease associations [14]. It is envisaged that the mechanisms of HLA expression variability can influence the strength of immune responses, for example, higher expression resulting in robust immune surveillance—mediated through larger numbers of HLA molecules for presenting foreign or tumor-specific peptides. On the contrary, higher expression of autoimmunity inducing HLA molecules are expected to elicit autoimmune responses. Further, polymorphic expression of HLA alleles represent another variable in determining outcome to disease beyond peptide binding alone, for example, through interacting with KIR and other molecules expressed on natural killer (NK) cells and other cells, thus differentially regulating immune responses. Hence, it is conceivable that identifying HLA allele-specific expression correlates could provide novel pathophysiological mechanisms in diseases. Moreover, delineating the underlying mechanisms responsible for expression variation may provide an opportunity to modify their expression therapeutically. In this context, select disease associations identified in GWAS have subsequently been observed dependent on HLA expression, for example, HIV with HLA-C expression and Hepatitis B Virus (HBV) with HLA-DP expression. In this section, we cover select examples of impact of such HLA expression variations on diseases.

### 17.4.1 Human leukocyte antigen-C expression and disease development

The functional relevance of HLA-C has been explored in details in HIV infection and disease progression. Besides interacting with KIR molecules and regulating NK cell activity, HLA-C is resistant to HIV viral protein “Nef” mediated inhibition, unlike HLA-A and HLA-B molecules, thus priming cytotoxic T-cell responses. Recently, studies have shown that alleles of this locus are expressed differentially and display a continuum of expression [15]. The HLA-C alleles with higher expression have been reported to be protective in controlling



viremia, when cohorts of viremic controllers (viral load (VL) <2000 copies/mL) and noncontrollers (VL >10,000 copies/mL) were analyzed. Similar correlation of HLA-C expression was observed with progression rates from seroconversion to low (<200) CD4 counts/ $\mu$ L. GWAS analysis have revealed an important variant at -35 kb (rs9264942), which lies in close proximity to HLA-C, which regulates its expression and efficiently controls HIV viral load [16]. It has been demonstrated that the -35 kb T allele is a low expressor as compared to the C allele. The high expressor -35 kb C alleles (include HLA-C\*08:01/02/03, \*12:02/03, \*01:02, \*02:02, \*06:02, \*05:01, and \*14:02) confer lower odds ratio (OR). Contrarily, the low expressor -35 kb T alleles (\*03:03/04, \*04:01, \*07:01/02, \*15:02/05/06, \*16:01/02/04, and \*17:01) contributes to higher OR for the development of AIDS [17,18]. Further, this enhanced immune surveillance through higher HLA-C expression is protective in HIV/AIDS, nevertheless, it is conceivable to hypothesize that such a higher expression-linked manifestation can increase the risk of autoimmune disease. To this end the evaluation in this direction revealed a strong correlation of HLA-C expression with the risk of Crohn's disease but not with ulcerative colitis in the case of control studies [19].

#### 17.4.2 Human leukocyte antigen-DP expression and increased risk of chronic HBV infection

Variations in HLA-DP region have been reported to be important determinants in chronic HBV infection as reported in a recent genome-wide study on Asian patients [20]. Similarly, a variant rs9277534 in the 3'UTR of HLA-DPB1 was found to be strongly associated in European and Africans [21]. This variant also correlated with levels of HLA-DPB1 at the transcriptional and HLA-DP cell-surface expression levels. Though the mechanisms of polymorphic expression are largely unknown, higher cell-surface expression of HLA-DP has been shown to increase the risk of chronic HBV infection.

#### 17.4.3 Human leukocyte antigen expression correlates in autoimmune diseases

Some examples of autoimmune diseases, for example, vitiligo and SLE are in direct relevance to this section. In autoimmune vitiligo, melanocytes are targeted by the autoreactive T cells. Though several significantly associated loci have been observed through GWAS analysis, SNPs in the XL-9 region (enhancer intergenic region between HLA-DR and DQ) were associated with higher HLA-DR and DQ expression and higher IFN- $\gamma$  and IL-1 $\beta$ , supporting their clinical relevance. Similar findings have been reported in SLE, which is another autoimmune disease characterized by the development of autoantibodies and loss of tolerance to self-antigens. Among various significant associations, XL-9 polymorphisms were observed, which could increase the expression of risk alleles by nearly fourfold [22,23].

#### 17.4.4 Low versus high expression mismatches in transplantation

It has been reported that low expression mismatches show improved tolerance as compared to high expression mismatches, for example, C\*03:03/C\*03:04 mismatches [24,25]. However, this mismatch in combination with mismatches in other loci could significantly influence the outcome as observed for mismatches in low-expression HLA loci—DRB3/4/5, DQ, and DP [26]. Considering the previously stated gradient of expression for various HLA-C alleles, a report suggests that the expression level of patient's mismatched HLA-C alleles could reasonably predict the risk of graft-versus-host disease (GVHD) [25]. This risk of acute GVHD is directly linked with the increased expression of HLA-C, which can also differentiate high-risk antigen recognition site (ARS) mismatches (high-expression residue 116 mismatch) from low-risk ARS mismatches. This is in direct context of a model of alloreactivity that requires the quantity of expressed HLA as well as ARS to define immunogenicity and tolerance to the graft. Similarly, the correlation of HLA-C expression was observed with two broad categories of KIR-binding epitopes, that is, C1 and C2, suggesting cumulative synergism of expression, HLA ligand/KIR receptor interactions, and ARS (residue 116) in immune responses in hematopoietic stem-cell transplantation outcome, particularly the development of GVHD. Further, mismatching against a high DPB1 expression patient resulted in significantly higher risk of GVHD in comparison to those carrying low-expression alleles. A stronger correlation was observed in cumulative analysis of HLA-DP expression along with the exon-2-defined epitopes, thus supporting the model of alloreactivity [26]. On the other hand, multiple mismatches at the low-expression HLA-DRB3, 4, and 5 genes have been shown to associate with adverse outcomes of the Hematopoietic stem cell transplantation (HSCT) [27].

### 17.4.5 Human leukocyte antigen class I expression loss/downregulation in tumor immune escape

HLA class I loss or downregulation is one of the most important immune escape mechanisms adopted by tumor cells. Various research groups have analyzed different HLA class I phenotypic aberrations in human tumors [28]. This loss of HLA on cell surface could either be reversible/soft or irreversible/hard and attributed to alterations in HLA,  $\beta 2$  microglobulin, or IFN genes. Incidentally, the efficacy of recent immunotherapeutic strategies also depends on level of expression of HLA class I on the target tumor cells. In case of reversible HLA loss, immunotherapy may upregulate expression of various cytokines in the tumor microenvironment, which can further repair the altered HLA expression. On the contrary, if the loss of HLA is irreversible/hard at the molecular level, the tumor cells are likely to remain HLA class I negative and compromise efficacy of any immune modulation. In recent years, accumulating evidence in this domain highlights the urgent need to unravel the molecular mechanisms involved in tumor-associated HLA loss, thus paving the way toward designing strategies to revert such aberrations.

### 17.4.6 Mechanisms underlying allele-specific human leukocyte antigen expression

Exact mechanisms underlying differential HLA allele-specific expression are largely unknown, and there is limited literature in this area [14]. Haplotype-specific expression is another dimension to this complexity of polymorphic expression. For example, HLA-DR15 haplotype is associated with significantly lower HLA-DRB1 expression and higher HLA-DQA1 and DQB1 expression than HLA-DR3 haplotypes [29]. Transcriptomic analyses representing three haplotypes, namely, HLA-A1-B8-Cw7-DR3, HLA-A3-B7-Cw7-DR15, and HLA-A26-B18-Cw5-DR3 revealed a high-resolution map uncovering several genes as signatures of interhaplotypic differences [30]. Among these, zinc finger proteins attributed to the most significant differences suggesting the role of transcriptional regulation and methylation in defining haplotype-specific differences in expression. Even though there are limited reports in this area, here we discuss here some mechanisms that might act either independently or synergistically to regulate/control the observed polymorphic HLA expression

1. *Methylation*: DNA methylation is an important epigenetic factor, which plays an important role in regulating gene expression. Aberrant methylation has been linked to several human diseases. As observed for HLA-C, a continuum of allele-specific expression has also been observed for HLA-A lineages by quantitative polymerase chain reaction (qPCR) [31]. However, in contrast to HLA-C, no polymorphisms in promoter and/or 3'UTR explain these expression differences. Further, both HLA-A and HLA-C share binding site for miR148a, this site is nonpolymorphic in former. On the other hand, epigenetic methylation patterns mediated downregulation. For example, representative low-expression HLA-A alleles (HLA-A\*03) are less methylated as compared to high-expression HLA-A alleles (A\*24) as revealed through bisulfite sequencing [31]. On the contrary, no significant influence of methylation was observed on the allele-specific expression of HLA-B and HLA-C alleles.
2. *Posttranslational influence*: In contrast to HLA-A and HLA-C, qPCR did not reveal any variation in allele-specific expression of HLA-B alleles at the transcription level. Similar results were observed by the RNA sequencing in two populations [32]. Further, HLA-B allele lineages lack the binding site for miR148a. However, there is some evidence of allele-specific differential expression of HLA-B lineages at the cell-surface level using antibodies for Bw4 or Bw6 epitopes. This could be attributed to posttranslational mechanisms, namely, assembly and export. Further complexity is observed in cell-specific differences in observed allele-specific expression, for example, the pattern of gradient expression for HLA-B differs in monocytes from lymphocytes.
3. *miRNA binding*: Variable expression of HLA-C is regulated by allele-specific polymorphisms. For example, an insertion of a single-nucleotide G at 263 (rs67384697-G) in the 3'UTR region of HLA-C alleles results in the formation of binding site for miRNA (miR148a), which leads to negative regulation and degradation of such alleles. The miR148a binding site is polymorphic in HLA-C loci only, although in HLA-A, it is nonpolymorphic while it is absent in HLA-B alleles [19]. This complexity is further enhanced by the polymorphism in miR148a gene (on chromosome 7) causing either high or lower expression of this miRNA.
4. *Enhancer polymorphisms-mediated expression variability*: Transcriptional enhancer XL-9 (within 130 kb intergenic region) between HLA-DRB1 and DQA1 is known to be involved in chromatin modeling [33]. HLA-DR-DQ expression is affected by variations in IRF4 and CTCF binding sites residing within XL-9. It is presumable that to certain extent haplotype-specific expression could be attributable to variation in this enhancer region.

Overall there is a need to unravel further beyond this abovementioned molecular complexity of HLA allele-specific expression patterns, particularly in the context of the HLA association and aberrations in various human diseases.

## 17.5 Human leukocyte antigen and organ transplantation

Early attempts at transplanting human organs beginning with the renal allografts by Soviet surgeon Prof. Yuri Yuriyevich Voronoy in 1933, by and large were unsuccessful until the introduction of immunosuppressive drugs in early 1960s, and realization about the involvement of immunological factors followed by the discovery of human HLA system. Simonsen [34] and Dempster [35] from independent studies in dogs concluded that an immunological mechanism was responsible for allograft failure and both considered that a humoral mechanism of rejection was likely. Soon, it was realized that for transplantation to succeed as a realistic form of renal replacement therapy, it was important to overcome the immunologic barrier. Matching donor and recipient pairs for histocompatibility antigens could not only suppress alloimmunity but also help in lowering the need for immunosuppressive drug therapy. One of the major landmarks in our understanding of the biological importance of human MHC has been the discovery of the molecular structure of HLA by the group of Jack Strominger, Don Wiley, and Pamela Bjorkman describing for the first time the crystal structure of HLA class I molecule in 1987 [36,37]. This particular discovery unraveled the mystery behind the mechanism of antigen presentation and allorecognition and further explained the phenomenon of MHC restriction. In recent years, advances in transplant immunobiology and the use of therapeutic strategies primarily directed against the T cell have led to a major shift in our thinking from considering allograft rejection being primarily a T cell-mediated process to the realization that alloantibodies especially those directed against specific HLA antigens of the donor impose a considerable barrier to both short-term as well as long-term allograft function and survival. Although the impact of HLA antibodies present at pretransplant stage has been noticed and understood as far back as 45 years ago [38], knowledge surrounding the importance/relevance of their de novo development has only been realized in more recent years [39,40]. Of late, an important role of non-HLA incompatibility between donor and recipient in influencing allograft immunity is increasingly being recognized due to ever growing studies reporting graft rejection and/or failure in renal transplants from HLA-matched donors [41].

### 17.5.1 Allorecognition

The term “allorecognition” refers to T-cell recognition of genetically encoded polymorphisms between members of the same species. The transplantation of tissues between genetically disparate individuals of the same species elicits a unique and complex immune response. The main target for the immune response to transplanted tissue is the MHC molecules expressed on donor cells, which are recognized as “nonself” to the recipient, generating an alloimmune response, and these remain a substantial barrier to solid organ and hematopoietic stem-cell transplantation. There are at least two distinct but not mutually exclusive pathways by which T lymphocytes recognize donor alloantigens, namely, the direct and indirect pathways of the alloresponse. More recently, a third mechanism termed semidirect allorecognition has also been described [42].

In the *direct pathway*, recipient T cells recognize intact donor MHC molecules on the surface of donor or allogeneic APCs. This pathway has long been known to be a potent means of eliciting acute allograft rejection, which is partly owing to the uniquely high frequency of T cells with direct allospecificity. The *indirect pathway*, on the other hand, involves the recognition of donor antigens that have been processed and presented in peptide form in the context of self-MHC on the host's own APCs. While APCs of the donor, within a transplanted organ, get gradually eliminated with passage of time, the recipient APCs constantly traffic through the graft. Upon migration to local lymph nodes, these cells are able to phagocytose, process, and present soluble alloantigen to the cytotoxic T cells. Studies carried out in the animal model systems as well as clinical settings have indicated that indirect allorecognition promotes chronic rejection. Further, evidence has been presented to suggest that recipient-derived endothelial cells can also play a role in the indirect pathway of allorecognition. Indeed, endothelium of the transplanted organs can progressively be replaced by recipient cells, which can also present alloantigens to the host immune cells via the indirect pathway [43]. In addition to the fact mentioned previously, a *semidirect pathway* has also been described. In this pathway the recipient APCs acquire and present intact donor

MHC class I molecules to stimulate CD8 + T cells via the direct pathway and simultaneously present internalized and processed donor MHC molecules as peptides to CD4 + T cells with indirect antidonor allospecificity.

### 17.5.2 Classification of rejection

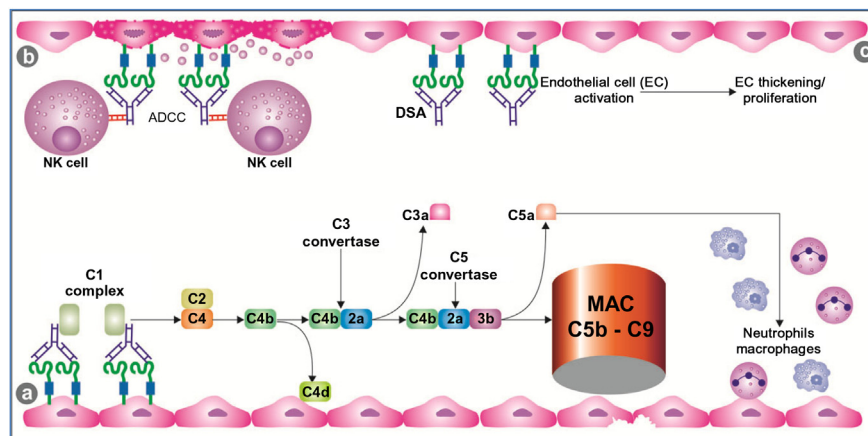
Based on the time of occurrence, allograft rejections have been categorized as hyperacute (occurring within minutes to hours), acute (days to months), and chronic rejection (months to years). Similarly based on the nature of underlying immunopathological mechanism, acute rejections have further been grouped into cellular and humoral rejection. It has long been recognized that the presence of preformed donor-specific antibodies (DSAs) in high titers in kidney recipients is associated with hyperacute rejection. Presently, the incidence of hyperacute rejection has dropped significantly and is extremely rare due to the universal adoption of pretransplantation complement-dependent cytotoxicity (CDC) crossmatching procedure. Nevertheless, preexisting antibodies or those developing de novo following renal transplantation are an important cause for acute as well as chronic rejection, particularly if these are donor specific [44–46].

### 17.5.3 Antibody-mediated rejection

Antibody formation is initiated by the uptake and processing of foreign antigens by the APCs that present them to the CD4 + helper T cells leading eventually to their clonal expansion and activation of B cells that secrete IgG through the generation of plasma cells. In the process, memory B cells get generated. The quiescent memory B cells and the long-lived plasma cells (LLPCs) are the two sources from which DSA is derived. LLPCs constitutively secrete antibodies and do not mobilize upon alloantigen reexposure unlike that of quiescent memory B cells. The latter accounts for the activation and generation of de novo DSAs upon alloantigen reexposure [47]. Mechanism by which antibodies may elicit damage to the graft has been the subject of much debate over the years. Experimental studies have identified three pathways, each of which are associated with the antibody-mediated rejection (AMR). These include complement activation, NK cell-mediated antibody-dependent cell-mediated cytotoxicity, and direct activation of the endothelium with antibodies (Fig. 17.3), which have been discussed in details in the following sections.

### 17.5.4 Complement activation

Alloantibodies developed as a result of humoral response to allograft are able to bind both to HLA and/or non-HLA targets expressed on endothelial cells of the graft microvasculature. Accordingly, the antigen-complexed



**FIGURE 17.3** Hypothesized mechanisms underlying the development of AMR. (A) The classical complement cascade being activated by antigen–antibody complex through C1qrs complex and ultimately leading to the formation of MAC and C4d, an inactive by product. The latter is deposited on endothelial cells. (B) NK cell-mediated ADCC is shown in upper left half (see text for explanation). (C) Direct activation of endothelium by HLA antibodies leading to cellular proliferation is shown on right half of the figure. ADCC, Antibody-dependent cell-mediated cytotoxicity; AMR, antibody-mediated rejection; HLA, human leukocyte antigen; MAC, membrane attack complex; NK, natural killer. Source: Mehra NK, Baranwal AK. Clinical and immunological relevance of antibodies in solid organ transplantation. *Int J Immunogenet* 2016;43 (6):351–68, reproduced with permission from John Wiley & Sons.



antibody molecules bind to the multivalent C1q molecule and trigger the classical complement pathway, which through a cascade of events leads to the production of C3 and C5 convertases and subsequent formation of MAC. The latter forms transmembrane pores that lead to the disruption of plasma membrane and lysis of the target cell. Complement components such as C3a and C5a act as chemoattractants for neutrophils and macrophages. The latter also acts as an anaphylatoxin and leads to the formation of edema due to the release of histamine from mast cells. Moreover, prostaglandin E2 released from macrophages due to the action of C3a causes vasospasm that further adds to the injury. Complement-mediated injury is particularly relevant to hyperacute and acute rejections. Antibody subclass and isotypes define effector functions of the molecule, including complement fixation. In humans, both the IgM and IgG subclasses, IgG1 and IgG3, are effective activators of the complement.

### 17.5.5 Antibody-dependent cell-mediated cytotoxicity

Here the NK cells through their Fc $\gamma$ RIII receptor (CD16) are able to bind the constant Fc region of the antibody molecule bound to HLA and/or non-HLA targets on the endothelium. This cross-linking of Fc receptors on NK cells to IgG antibody-coated cells triggers perforin–granzyme-mediated cytotoxicity of target cells.

### 17.5.6 Direct activation of endothelium

HLA antibodies, besides activating the complement cascade, can also directly activate the endothelium especially cellular proliferation, which is central to the pathogenesis of chronic rejection. Unlike the activation of complement cascade, which predominantly causes the lysis of endothelial cells, direct activation of endothelium by antibodies is dominated by sublytic endothelial injury and structural modification of microcirculation. Morphologically they manifest as endothelial cell proliferation and/or thickening, duplication and multilayering of basement membrane of glomeruli and peritubular capillaries.

### 17.5.7 Cellular rejection

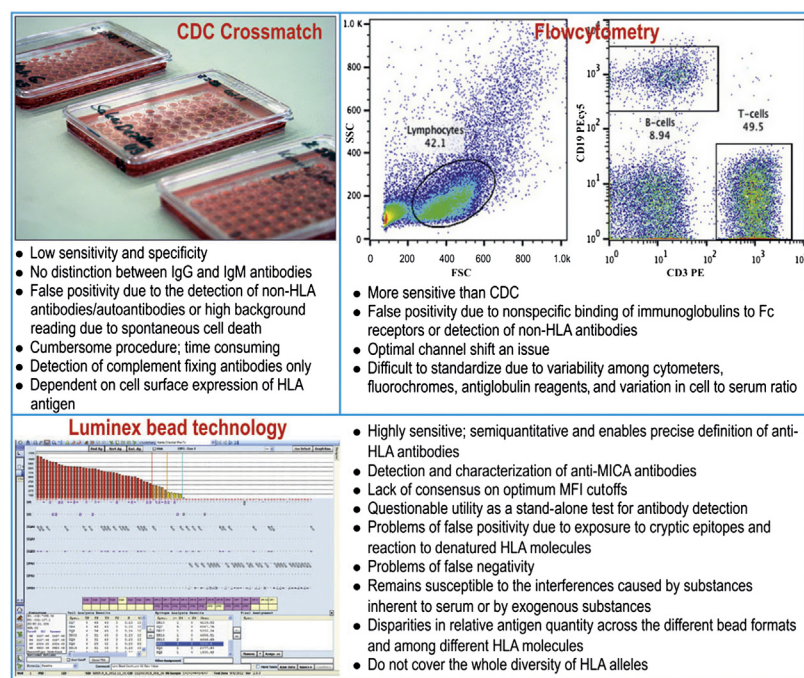
Alloantigen recognition is the primary event in the cascade of alloreactive T-cell activation and proliferation leading to rejection of the transplanted tissue/organ. Although the precise mechanism by which T cells mediate graft injury is not certain, two models have been proposed: (1) cell-mediated cytotoxicity of parenchymal cells (tubular and endothelial) and (2) through the effects of local cytokine release, analogous to a delayed-type hypersensitivity reaction. Cytokines may act directly on parenchymal cells or indirectly through effects on the endothelium and vascular supply.

## 17.6 Human leukocyte antigen–antibody-detection techniques

Antibody-detection techniques have evolved significantly providing an ever increasing degree of sensitivity and specificity, from the good old conventional cell-based assays such as the CDC and flow crossmatch to the currently advanced solid-phase systems based on the Luminex platform [45]. Fig. 17.4 provides a realistic comparison between the three main techniques for the detection of anti-HLA antibodies. Studies carried by us have revealed that DSAs directed against HLA antigens detected by CDC are a contraindication for transplantation, whereas those detected by assays other than the CDC represent varying degrees of risk depending on the level of antibody titer. A comparison of the three techniques revealed that Luminex-based DSA mean fluorescence intensity (MFI) values correlated with the CDC and flow crossmatch (FCXM) positivity only when the mean MFI reached a cutoff of 3000 and 7000, respectively [48].

Besides being the most sensitive assay, antibody screening on Luminex platform is highly accurate in defining donor-specific HLA antibodies in broadly reactive alloantisera. The test is of great value to guide donor selection and kidney-paired exchange programs through virtual crossmatching and for instituting desensitization measures at pretransplant stage and for posttransplant monitoring of DSAs. Currently, there is a lack of consensus with regard to the optimum MFI “cutoffs” for classifying antibodies as significant. In a bid to improve sensitivity and specificity of antibody detection, methods such as flow cytometry and solid-phase assays compromise discrimination between complement fixing antibodies from the noncomplement fixing types. Development of single-antigen bead (SAB) assay has been a remarkable breakthrough since it provided a system to detect DSAs against single HLA allele with unprecedented sensitivity and specificity. Nevertheless, data from several studies





**FIGURE 17.4** A summarized view of the cell-based (CDC crossmatch and flow cytometry) and solid phase-based assays (Luminex-based technology) for the detection of anti-HLA antibodies, highlighting main advantages and disadvantages. CDC, Complement-dependent cytotoxicity; HLA, human leukocyte antigen. Source: Mehra NK, Baranwal AK. *Clinical and immunological relevance of antibodies in solid organ transplantation*. *Int J Immunogenet* 2016;43(6):351–68, reproduced with permission from John Wiley & Sons.

have indicated that not all Luminex detectable antibodies may be clinically relevant and that new tools need to be developed to distinguish inert antibodies from the deleterious ones.

C1q assay that is essentially a modification of the solid-phase assay is able to detect complement binding antibodies, thus retaining the exquisite sensitivity and specificity of the Luminex platform. Similarly, the C3d assay, although very similar to C1q in methodology, does not utilize a recombinant complement product. Hence, this could be more accurate to predict in vivo complement activation. A recent review highlights the technical challenges and clinical relevance of SAB C1q and C3d assays as well as IgG subclass analysis of donor-specific anti-HLA antibodies [49].

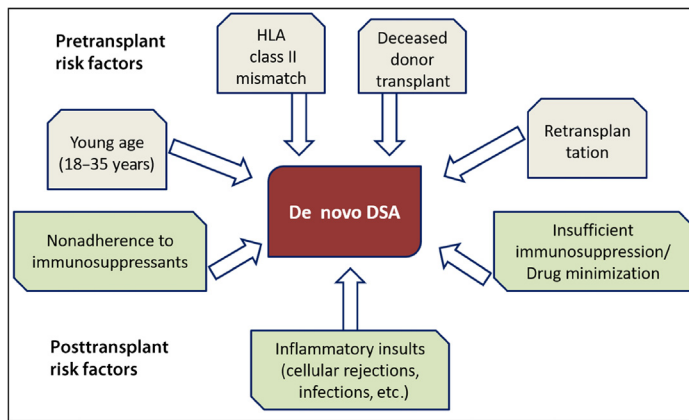
### 17.6.1 Relevance of anti–human leukocyte antigen antibodies

Technological advancements in antibody detection have provided increasingly better tools to understand their role in AMR and allograft loss. While the presence of pretransplant antibodies has been attributed to blood transfusion, pregnancy, and prior transplantation or sometimes to unknown causes, the potential risk factors for the development of de novo DSAs include younger age of the patient, HLA class II mismatch status between donor and the recipient, noncompliance with medications and inflammatory insults such as infections and/or cellular rejections [39,40]. Fig. 17.5 illustrates possible risk factors for the development of DSA both at pre- and posttransplant stages.

It is now clear that a combined presence of anti-HLA class I and II antibodies lead to more frequent rejection episodes as compared to patients having either of these antibodies [50,51]. The development of de novo DSA is one of the earliest and a well-established biomarker in posttransplant setting to predict allograft injury and loss. Such antibodies are associated with late acute and chronic AMR, and toward transplant glomerulopathy [40,52,53]. Evidence has been provided to suggest that antibodies are generally produced long before the observed elevation of serum creatinine or development of rejection episodes [39]. Further, performance of protocol allograft biopsy at the time of first detection of de novo DSA uncovers subclinical AMR disease [54]. This highlights the importance of systematic monitoring of posttransplant HLA-DSA for early diagnosis of AMR.

### 17.6.2 Role of non–human leukocyte antigen antibodies

Although antibodies specific to donor HLA antigens are responsible for a large part of the AMR, a fair number of these rejections are attributed to the presence of non-HLA group of antibodies. The most notable of the non-HLA antigen targets that have been shown to influence the graft outcome in solid organ transplantation include



**FIGURE 17.5** Possible risk factors for the development of donor-specific antibodies at pretransplant stage (highlighted in gray) and de novo antibody occurrence at posttransplant stages (highlighted in green).

MHC class I chain-related molecule A (MICA), angiotensin II type 1 receptor, vimentin, collagen, cardiac myosin, and K- $\alpha$ 1 tubulin. Of these, antibodies against MICA have been most promising and extensively studied. We and others have reported a strong association between the presence of MICA antibodies (at pre- or posttransplant stages) and adverse graft outcome [55,56]. Further, the presence of anti-MICA antibodies alone in the absence of anti-HLA antibodies could also cause not only acute and/or chronic rejection but in certain cases also the hyperacute rejection [57]. However, due to a lack of unanimity, this relationship has not been fully established as yet. There is also no consensus with regard to the optimum MFI “cutoffs” for classifying MICA antibodies as clinically relevant. The risk factors for development of MICA antibodies are mainly the same as that of HLA sensitization, that is, pregnancy, blood transfusion, and previous transplantation.

### 17.6.3 Preventive measures

Recent years have seen remarkable improvements not only toward effective immunosuppression protocols, but also on the readily available tools and biologicals for antibody minimization. This has made it possible to tackle the issue of early acute rejection to a large extent. However, late rejection and long-term graft survival is still a cause of concern, and evidence has been presented to suggest that host-mediated alloimmune responses and development of DSAs are the central issues [58]. Therefore prevention of antibody-mediated allograft damage is of prime importance. Such preventive measures should start before transplantation, beginning with avoidance of sensitizing events, continuing through peritransplantation measures to posttransplant monitoring.

Blood transfusion, pregnancy, and previous organ transplantation are the three well-documented events by which a patient gets sensitized to HLA antigens, and such patients are known to have adverse outcomes. Accordingly, it is important to avoid administration of blood and blood products as far as possible. Several studies have highlighted the importance of better HLA matching as it not only leads to better graft survival both for deceased and live donor transplants, but also associated with lower grades of sensitization, which is so very needed if a retransplant is required. Better HLA matching between donor–recipient pairs is also associated with a reduction in the incidence of de novo DSA development [39,59,60]. Of late, more precise matching for HLA epitopes based on HLA Matchmaker is being prophesied [61]. Further, HLA class I- and II-specific antibodies should be precisely characterized by solid-phase immunoassays, particularly for sensitized patients. To achieve this, high-resolution HLA testing of donor–recipient pairs is recommended. Ideally a minimum of two sera obtained at different time points should be tested to confirm the presence or absence of antibodies. Removal of DSA from the recipients’ circulation or to circumvent it is the only way forward for the timely transplantation of sensitized patients.

There are two options available for successful transplantation of such patients. First is their inclusion in special programs such as kidney-paired donation and through “acceptable mismatch” strategy. Both approaches involve desensitization. Alloantibodies may be removed from the patients’ circulation either by plasmapheresis or immunoadsorption and intravenous immunoglobulins. This can be combined with other therapeutic strategies such as depletion of B-lymphocytes by rituximab, or plasma cells by the use of bortezomib and complement inhibitor (C-5) eculizumab. Different centers have adopted different desensitization protocols. The use of individualized adequate and appropriate immunosuppression is one of the most important aspects of preventive strategies. Insufficient immunosuppression makes patients prone to cellular rejection, which in itself is a risk factor for the development of de novo DSA and subsequently to AMR and graft loss [39]. On the other hand, strong immunosuppression is

associated with the development of side effects such as infections and adverse cardiac events. Proinflammatory events such as infections and myocardial infarctions are also known to be associated with significant increase in both breadth and strength of HLA-specific antibodies. A systematic review by Butler et al. [62] have revealed that 23%–50% of recipients are noncompliant in the form of either intermittent omission of few drugs, avoidance of a particular drug, surreptitious dose reduction, or complete cessation of therapy. Incidence of poor medication adherence rises with increasing transplant duration. Noncompliance may lead to AMR and graft loss via the development of de novo DSA. The main strategies to prevent nonadherence include repeated individualized education of recipients, minimization of drug costs and side effects and regular clinical support addressing patient's concerns.

## 17.7 Human leukocyte antigen and blood transfusion

HLA alloimmunization mediates some of the serious complications such as febrile nonhemolytic transfusion reactions (FNHTR), immunological platelet refractoriness, transfusion-related acute lung injury (TRALI), and transfusion-associated GVHD (TA-GVHD) developing following the transfusion of blood and blood products. Although universal leucodepletion practiced by almost all centers has by far eliminated the main source of HLA alloimmunization, residual cellular components and platelets can induce the development of anti-HLA antibodies and alloreactive T cells.

It is now amply clear that FNHTR and immunological refractoriness to platelets are mediated mainly by the presence of antibodies in the recipient reacting against HLA class I molecules in the transfused products. Posttransfusion platelet increment (PI) is the most commonly used method to assess the effectiveness of platelet transfusion. A PI of  $>10 \times 10^9/L$  at 1- or 24-hours posttransfusion is considered an adequate response [63]. Platelet refractoriness with nonimmune etiology will typically have a normal response in platelet count within 1 hour of transfusion but will often return to baseline within 24 hours. Immunological refractoriness to platelets, however, typically show little-to-no increase in platelet count at 1-hour posttransfusion [64]. Immune refractoriness is mainly attributed to alloimmunization either against HLA and/or human platelet antigen (HPA) due to prior exposure from transfusion, pregnancy, or transplantation. HLA alloimmunization is more commonly implicated as a cause of immune refractoriness to platelets than HPA antibodies. Since platelets express only class I HLA molecules with expression of HLA-A and -B being much higher than HLA-C, anti-HLA antibodies directed against anti-HLA-A and -B antigens are significant in causing platelet refractoriness unlike that against anti-HLA-C antigens. The limited studies in this area have shown conflicting reports on the role of anti-HLA-C antibodies in the causation of platelet refractoriness [65–67]. This is the reason why matching for HLA-C alleles has not regularly been considered in the provision of HLA-selected platelets.

Variable approaches have been used to overcome the immunological refractoriness to platelets. These include (1) transfusion of crossmatch compatible platelets. This is generally useful in patients who are not highly sensitized but have the disadvantage of potential immunization against mismatched HLA antigens, hence not suited for long-term platelet transfusion support, (2) transfusion of virtually crossmatched platelets based on HLA antibody profile of the patient and not on the degree of HLA matching between patient and the donor. This is useful for short-term platelet transfusion support but compromises the efficacy of future transfusions because of the risk of immunizing the patient to the mismatched antigens, and (3) the third and the most logical approach is based on the transfusion of selected platelets based on the HLA phenotype status of the recipient and the antibody profile as determined through panel reactive antibody screening.

An analysis of results of the British "Serious Hazards of Transfusion" initiative as well as hemovigilance systems in many other countries have demonstrated TRALI as one of the most common causes of transfusion-associated major morbidity and death [68,69]. TRALI has been shown to be associated with the presence of antibodies against both HLA class I and II molecules, as well as neutrophil-specific antibodies in the plasma of both donors and recipients [70,71]. Majority of these reactions have been observed with multiparous female blood donors who have been sensitized to fetal antigens. Indeed, the use of plasma from male donors and nulliparous or multiparous female donors who are negative for leukocyte-directed antibodies has drastically reduced the incidence and severity of TRALI [72,73]. Similarly transfusions of leuko-reduced cellular blood products have led to the reduction in the prevalence of TRALI being caused by leukocyte antibodies present in the recipients.

TA-GVHD results from the transfusion of immunologically responsive lymphocytes into immunosuppressed recipients and is almost uniformly fatal [74]. It is seen in 0.1%–1% of all transfusions in susceptible recipients [75]. All cellular blood products that contain viable immunocompetent T cells, that is, red cells, platelet and

granulocyte concentrate have been implicated in TA-GVHD. In situations where the donor is homozygous for both HLA haplotypes while the recipient is heterozygous, the risk of TA-GVHD is significantly increased. Owing to the zero HLA mismatch in the direction of host versus graft, donor lymphocytes are not rejected as foreign. However, the HLA mismatch in the direction of graft versus host arising due to homozygous donor, T cells perceives host cells as foreign and react against them. Evidence of chimerism, that is, the presence of donor-derived cells, chromosomes, or DNA in the blood and/or affected tissues of the recipient is essential for the diagnosis of TA-GVHD.

## 17.8 Conclusions

Human MHC reflects several inherent features, which clearly distinguishes it from rest of the genome. The region continues to reveal significant associations with various diseases as observed through candidate gene-based approach and GWAS across worldwide populations. Although various mechanisms have been put forward to explain HLA and disease associations, definitive involvement of these genes themselves remains illusive. Besides the extraordinary polymorphism within HLA system, allele-specific expression on specific tissues puts an additional layer of complexity beyond the genomic diversity. However, the novel mechanisms regulating such expression variability and its importance in the context of disease development and/or immune aberrations are largely unknown.

Postgenomic era has witnessed remarkable progress toward identifying novel alleles at high-resolution level through advancements in high-throughput molecular technologies, namely, solid-phase assays, next-generation sequencing, and others. All these have led to a better understanding of the HLA-associated diseases, paving the way for developing robust diagnostic and therapeutic approaches. Role of HLA matching, antibody detection, and involvement of other hitherto unexplored factors are relevant in long-term acceptance of the organ and hematopoietic stem-cell graft. Indeed, chronic alloimmune response is the main driver of late allograft loss, and this will remain one of the major area of future research. Immunological stratification is required to design individualized immunosuppressive regimens, and this needs further research on genome-derived HLA and non-HLA factors.

## References

- [1] Mallal S, Nolan D, Witt C, et al. Association between presence of HLA-B\*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet* 2002;359(9308):727–32.
- [2] Hung SI, Chung WH, Liou LB, et al. HLA-B\*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol. *Proc Natl Acad Sci USA* 2005;102(11):4134–9.
- [3] Zhang FR, Liu H, Irwanto A, et al. HLA-B\*13:01 and the dapsone hypersensitivity syndrome. *N Engl J Med* 2013;369(17):1620–8.
- [4] Man CB, Kwan P, Baum L, et al. Association between HLA-B\*1502 allele and antiepileptic drug-induced cutaneous reactions in Han Chinese. *Epilepsia* 2007;48(5):1015–18.
- [5] Schutte RJ, Sun Y, Li D, Zhang F, Ostrov DA. Human leukocyte antigen associations in drug hypersensitivity reactions. *Clin Lab Med* 2018;38(4):669–77.
- [6] Hughes T, Adler A, Kelly JA, et al. Evidence for gene-gene epistatic interactions among susceptibility loci for systemic lupus erythematosus. *Arthritis Rheum* 2012;64(2):485–92.
- [7] Kirino Y, Bertsias G, Ishigatsubo Y, et al. Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLAB\*51 and ERAP1. *Nat Genet* 2013;45(2):202–7.
- [8] Cortes A, Pulit SL, Leo PJ, et al. Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. *Nat Commun* 2015;6:7146.
- [9] Castro-Santos P, Moro-García MA, Marcos-Fernández R, et al. ERAP1 and HLA-C interaction in inflammatory bowel disease in the Spanish population. *Innate Immun* 2017;23(5):476–81.
- [10] Martin MP, Naranbhai V, Shea PR, Qi Y, Ramsuran V, Vince N, et al. Killer cell immunoglobulin-like receptor 3DL1 variation modifies HLA-B\*57 protection against HIV-1. *J Clin Invest* 2018;128(5):1903–12.
- [11] Kennedy AE, Ozbek U, Dorak MT. What has GWAS done for HLA and disease associations? *Int J Immunogenet* 2017;44(5):195–211.
- [12] Busch R, De Riva A, Hadjinicolaou AV, Jiang W, Hou T, Mellins ED. On the perils of poor editing: regulation of peptide loading by HLA-DQ and H2-A molecules associated with celiac disease and type 1 diabetes. *Expert Rev Mol Med* 2012;14:e15.
- [13] Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 2009;373(9664):659–72.
- [14] Petersdorf EW, O'Huigin C. The MHC in the era of next-generation sequencing: implications for bridging structure with function. *Hum Immunol* 2019;80(1):67–78.
- [15] Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, et al. Influence of HLA-C expression level on HIV control. *Science* 2013;340:87–91.
- [16] Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007;317:944–7.



- [17] Kulpa DA, Collins KL. The emerging role of HLA-C in HIV-1 20 infection. *Immunology* 2011;134:116–22.
- [18] Thomas R, Apps R, Qi Y, Gao X, Male V, O'huigin C, et al. HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. *Nat Genet* 2009;41:1290–4.
- [19] Kulkarni S, Qi Y, O'huigin C, Pereyra F, Ramsuran V, McLaren P, et al. Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proc Natl Acad Sci USA* 2013;110(51):20705–10.
- [20] Kamatani Y, Wattanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, Hosono N, et al. A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat Genet* 2009;41:591–5.
- [21] Thomas R, Thio CL, Apps R, Qi Y, Gao X, Marti D, et al. A novel variant marking HLA-DP expression levels predicts recovery from hepatitis B virus infection. *J Virol* 2012;86:6979–85.
- [22] Raj P, Rai E, Song R, Khan S, Wakeland BE, Viswanathan K, et al. Regulatory polymorphisms modulate the expression of HLA class II molecules and promote autoimmunity. *Elife* 2016;5:e12089.
- [23] Cavalli G, Hayashi M, Jin Y, Yorgov D, Santorico SA, Holcomb C, et al. MHC class II super-enhancer increases surface expression of HLA-DQ and affects cytokine production in autoimmune vitiligo. *Proc Natl Acad Sci USA* 2016;113:1363–8.
- [24] Fernández-Viña MA, Wang T, Lee SJ, Haagenson M, Aljurf M, Askar M, et al. Identification of a permissible HLA mismatch in hematopoietic stem cell transplantation. *Blood* 2014;123:1270–8.
- [25] Petersdorf EW, Gooley TA, Malkki M, Bacigalupo AP, Cesbron A, Du E, et al. International Histocompatibility Working Group in Hematopoietic Cell Transplantation, HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation. *Blood* 2014;124:3996–4003.
- [26] Fernández-Viña MA, Klein JP, Haagenson M, Spellman SR, Anasetti C, Noreen H, et al. Multiple mismatches at the low expression HLA loci DP, DQ, and DRB3/4/5 associate with adverse outcomes in hematopoietic stem cell transplantation. *Blood* 2013;121:4603–10.
- [27] Petersdorf EW, Malkki M, O'huigin C, Carrington M, Gooley T, Haagenson MD, et al. High HLA-DP expression and graft-versus-host disease. *N Engl J Med* 2015;373:599–609.
- [28] Garrido F, Perea F, Bernal M, Sánchez-Palencia A, Aptsiauri N, Ruiz-Cabello F. The Escape of Cancer from T Cell-Mediated Immune Surveillance: HLA Class I Loss and Tumor Tissue Architecture. *Vaccines (Basel)* 2017;5(1):7.
- [29] Pelikan RC, Kelly JA, Fu Y, Lareau CA, Tessneer KL, Wiley GB, et al. Enhancer histone-QTLs are enriched on autoimmune risk haplotypes and influence gene expression within chromatin networks. *Nat Commun* 2018;9:2905.
- [30] Vandiedonck C, Taylor MS, Lockstone HE, Plant K, Taylor JM, Durrant C, et al. Pervasive haplotypic variation in the spliceo-transcriptome of the human major histocompatibility complex. *Genome Res* 2011;21:1042–54.
- [31] Ramsuran V, Kulkarni S, O'huigin C, Yuki Y, Augusto DG, Gao X, et al. Epigenetic regulation of differential HLA-A allelic expression levels. *Hum Mol Genet* 2015;24:4268–75.
- [32] Ramsuran V, Hernández-Sánchez PG, O'huigin C, Sharma G, Spence N, Augusto DG, et al. Sequence and phylogenetic analysis of the untranslated promoter regions for HLA class I genes. *J Immunol* 2017;198:2320–9.
- [33] Majumder P, Gomez JA, Boss JM. The human major histocompatibility complex class II HLA-DRB1 and HLA-DQA1 genes are separated by a CTCF-binding enhancer-blocking element. *J Biol Chem* 2006;281:18435–43.
- [34] Simonsen M. Biological incompatibility in kidney transplantation in dogs: serological investigations. *Acta Pathol Microbiol Scand* 1953;32(1):35–84.
- [35] Dempster WJ. Kidney homo transplantation. *Br J Surg* 1953;40:447–65.
- [36] Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 1987;329:506–12.
- [37] Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 1987;329:512–18.
- [38] Patel R, Terasaki PI. Significance of a positive crossmatch test in kidney transplantation. *N Engl J Med* 1969;280:735–9.
- [39] Wiebe C, Gibson IW, Blydt-Hansen TD, Karpinski M, Ho J, Storsley LJ, et al. International Histocompatibility Working Group in Hematopoietic Cell Transplantation, HLA-C expression levels define permissible mismatches in hematopoietic cell transplantation. *12:1157–67*.
- [40] Everly MJ, Rebellato LM, Haisch CE, Ozawa M, Parker K, Briley KP, et al. Incidence and impact of de novo donor-specific alloantibody in primary renal allografts. *Transplantation* 2013;95(3):410–17.
- [41] Reindl-Schwaighofer R, Heinzl A, Kainz A, van Setten J, Jelencsics K, et al. Contribution of non-HLA incompatibility between donor and recipient to kidney allograft survival: genome-wide analysis in a prospective cohort. *Lancet* 2019;393(10174):910–17.
- [42] Herrera OB, Golshayan D, Tibbott R, et al. A novel pathway of alloantigen presentation by dendritic cells. *J Immunol* 2004;173:4828–37.
- [43] Kapessidou Y, Habran C, Buonocore S, Flamand V, Barvais L, Goldman M, et al. The replacement of graft endothelium by recipient-type cells conditions allograft rejection mediated by indirect pathway CD4+ T cells. *Transplantation* 2006;82:582–91.
- [44] Terasaki PI, Cai J. Human leukocyte antigen antibodies and chronic rejection: from association to causation. *Transplantation* 2008;86:377–83.
- [45] Mehra NK, Baranwal AK. Clinical and immunological relevance of antibodies in solid organ transplantation. *Int J Immunogenet* 2016;43(6):351–68.
- [46] Mehra NK, Baranwal AK, Tait BD. Editoriol: Antibody Repertoire and Graft Outcome following Solid Organ Transplantation. *Front Immunol* 2017;8:648.
- [47] Chong AS, Sciammas R. Memory B cells in transplantation. *Transplantation* 2015;99(1):21–8.
- [48] Baranwal AK, Bhat DK, Goswami S, Kaur J, Kaur G, Agarwal SK, et al. Comparative analysis of Luminex based donor specific antibody (DSA) MFI values with CDC and Flow crossmatch results in live donor renal transplantation. *Indian J Med Res* 2017;145:222–8.
- [49] Karahan GE, Claas FHJ, Heidt S. Technical challenges and clinical relevance of single antigen bead C1q/C3d testing and IgG subclass analysis of human leukocyte antigen antibodies. *Transpl Int* 2018;31(11):1189–97.



- [50] Panigrahi A, Gupta N, Siddiqui JA, Margoob A, Bhowmik D, Guleria S, et al. Monitoring of anti-HLA and anti-Major histocompatibility complex class I related chain A antibodies in living related renal donor transplantation. *Transplant Proc* 2007;39(3):759–60.
- [51] Dieplinger G, Ditt V, Arns W, Huppertz A, Kisner T, et al. Impact of de novo donor-specific HLA antibodies detected by Luminex solid-phase assay after transplantation in a group of 88 consecutive living-donor renal transplantations. *Transpl Int* 2014;27(1):60–8.
- [52] Susal C, Wettstein D, Dohler B, Morath C, Ruhenstroth A, Scherer S, et al. Collaborative transplant study report: association of kidney graft loss with de novo produced donor specific and non-donor-specific HLA antibodies detected by single antigen testing. *Transplantation* 2015;99:1976–80.
- [53] Zhang R. Donor specific antibodies in kidney transplant recipients. *Clin J Am Soc Nephrol* 2018;13(1):182–92.
- [54] Loupy A, Vernerey D, Tinel C, Aubert O, Duong van Huyen JP, et al. Subclinical rejection phenotypes at 1 year post-transplant and outcome of kidney allografts. *J Am Soc Nephrol* 2015;26:1721–31.
- [55] Panigrahi A, Gupta N, Siddiqui JA, Margoob A, Bhowmik D, Guleria S, et al. Post-transplant development of MICA and anti-HLA antibodies is associated with acute rejection episodes and renal allograft loss. *Hum Immunol* 2007;68:362–7.
- [56] Sanchez-Zapardiel E, Castro-Panete MJ, Castilloroma M, et al. Harmful effect of preformed anti-MICA antibodies on renal allograft evolution in early post transplantation period. *Transplantation* 2013;96:70–8.
- [57] Mehra NK, Siddiqui J, Baranwal A, Goswami S, Kaur G. Clinical relevance of antibody development in renal transplantation. *Ann N Y Acad Sci* 2013;1283:30–42.
- [58] Lee PC, Ozawa M, Hung CJ, Lin YJ, Chang SS, Chou TC. Eighteen-year follow-up of a retrospective study of HLA antibody on kidney graft survival. *Transplant Proc* 2009;41(1):121–3.
- [59] Opelz G, Dohler B. Effect of human leukocyte antigen compatibility on kidney graft survival: comparative analysis of two decades. *Transplantation* 2007;84(2):137–43.
- [60] Baranwal AK, Singh YP, Mehra NK. Should we revisit HLA matching to improve long term graft outcomes? *Curr Transpl Rep* 2018;. Available from: <https://doi.org/10.1007/s40472-018-0201-1>.
- [61] Duquesnoy RJ. HLA epitope-based matching for transplantation. *Curr Opin Transpl Immunol* 2014;31:1–6.
- [62] Butler JA, Roderick P, Mullee M, Mason JC, Peveler RC. Frequency and impact of nonadherence to immunosuppressants after renal transplantation: a systematic review. *Transplantation* 2004;77:769–76.
- [63] Hod E, Schwartz J. Platelet transfusion refractoriness. *Br J Haematol* 2008;142:348–60.
- [64] Schmidt AE, Refaai MA, Coppage M. HLA-mediated platelet refractoriness. *Am J Clin Path* 2018;XX:1–11.
- [65] Datema G, Stein S, Eijssink C, et al. HLA-C expression on platelets: studies with an HLA-Cw1-specific human monoclonal antibody. *Vox Sang* 2000;79:108–11.
- [66] Saito S, Ota S, Seshimo H, et al. Platelet transfusion refractoriness caused by a mismatch in HLA-C antigens. *Transfusion* 2002;42:302–8.
- [67] Stanworth SJ, Hudson CL, Estcourt LJ, et al. TOPPS Study Investigators Risk of bleeding and use of platelet transfusions in patients with hematologic malignancies: recurrent event analysis. *Haematologica* 2015;100:740–7.
- [68] Williamson LM, Lowe S, Love EM, et al. Serious hazards of transfusion (SHOT) initiative: analysis of the first two annual reports. *Br Med J* 1999;319:16–19.
- [69] Keller-Stanislawski B, Reil A, Günay S, et al. Frequency and severity of transfusion-related acute lung injury – German haemovigilance data (2006–2007). *Vox Sang* 2010;98:70–7.
- [70] Silliman CC, Boshkov LK, Mehdizadehkashi Z, et al. Transfusion-related acute lung injury: epidemiology and a prospective analysis of etiologic factors. *Blood* 2003;101:454–62.
- [71] Reil A, Keller-Stanislawski B, Guenay S, et al. Specificities of leukocyte alloantibodies in transfusion-related acute lung injury and results of leukocyte antibody screening of blood donors. *Vox Sang* 2008;95:313–17.
- [72] Chapman CE, Stainsby D, Jones H, et al. Ten years of hemovigilance reports of transfusion-related acute lung injury in the United Kingdom and the impact of preferential use of male donor plasma. *Transfusion* 2009;49:440–52.
- [73] Eder AF, Herron RM, Strupp A, et al. Effective reduction of transfusion related acute lung injury risk with male predominant plasma strategy in the American Red Cross (2006–2008). *Transfusion* 2010;50(8):1732–42.
- [74] Dwyre DM, Holland PV. Transfusion-associated graft-versus host disease. *Vox Sang* 2008;95:85–93.
- [75] Gokhale SG, Gokhale SS. Transfusion-associated graft versus host disease (TAGVHD)—with reference to neonatal period. *J Matern Fetal Neonatal Med* 2015;28(6):700–4.

## Further reading

- Flesch BK, Neppert J. Transfusion related acute lung injury caused by human leucocyte antigen class II antibody. *Br J Haematol* 2002;116:673–6.
- Hamilton D, Reid WA. Yu Yu Voronoy and the first human kidney allograft. *Surg Gynecol Obstet* 1984;159:289–94.
- Liefeldt L, Brakemeier S, Glander P, Waiser J, Lachmann N, Schönemann C, et al. Donor-specific HLA antibodies in a cohort comparing everolimus with cyclosporine after kidney transplantation. *Am J Transplant* 2012;12(5):1192–8.
- Mehra NK, Kaur G. The HLA complex in biology and medicine. A resource book. New Delhi: Jaypee Brothers Medical Publishers; 2010.
- Ruiz P, Maldonado P, Hidalgo Y, Gleisner A, Sauma D, Silva C, et al. *Clin Dev Immunol* 2013;2013:210506.
- Willicombe M, Brookes P, Sergeant R, Santos-Nunez E, Steggar C, Galliford J, et al. De novo DQ donor-specific antibodies are associated with a significant risk of antibody-mediated rejection and transplant glomerulopathy. *Transplantation* 2012;94(2):172–7.



# Disorders of abnormal hemoglobin

*Reena Das and Prashant Sharma*

Department of Hematology, Postgraduate Institute of Medical Education and Research, Chandigarh, India

## 18.1 Introduction

Disorders of hemoglobin (Hb) include quantitative and qualitative defects in the synthesis of the globin chains and, rarely, heme. The former comprise the commonest monogenic disorders worldwide, and their studies span diverse disciplines including hematology, biochemistry, genetics, anthropology, and public health. Clinical and laboratory research on Hb has formed the basis of much of our current knowledge of human physiology at the molecular level, and also the translation of these genetic insights has often resulted in improved patient care. This chapter deals sequentially with the structure and function of normal Hb, the thalassemias, and qualitative defects including variant Hbs and those with structural instability or high/low affinity for oxygen.

## 18.2 The hemoglobin molecule: normal structure and function

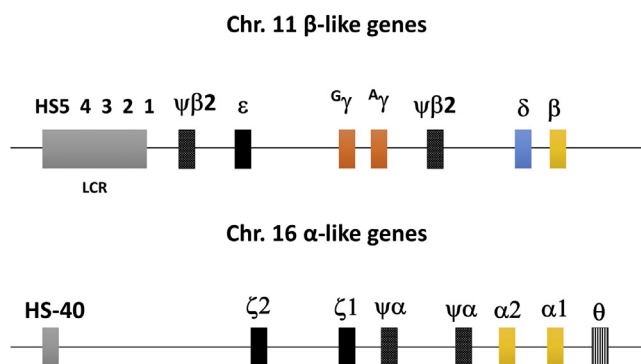
The structure of Hb was determined using high-resolution X-ray crystallography [1]. It is packaged in erythrocytes at ~31–35 g% and consists of heme (iron and porphyrin) and globin (protein) moieties. The tetrameric molecule comprises two  $\alpha/\alpha$ -like (141 amino acid) and two  $\beta/\beta$ -like (146 amino acid) globin chains. The globin chains fold into three-dimensional tertiary structures with a globin pocket with charged amino acids on the hydrophilic surface and uncharged amino acids in the hydrophobic interior. Each globin polypeptide has eight  $\alpha$ -helical segments (A–H), a heme group that binds to oxygen, ferroprotoporphyrin IX and is coordinated by a histidine at the eighth residue of the F helix (His-87 of  $\alpha$ -chain and His-92 of  $\beta$  chain), which is wedged into the pocket of each globin chain [2].

### 18.2.1 Globin gene clusters: structure and its regulation

The two globin gene families localize to different chromosomes wherein seven genes are present in each gene cluster, including functional and pseudogenes (Fig. 18.1). The genes are transcriptionally expressed as 5'–3' orientation from embryogenesis. The  $\alpha$ -globin gene cluster is located in the telomeric region of chromosome 16 (16p13.3), and the functional genes are embryonic zeta ( $\zeta$ ) and two adult duplicated  $\alpha 2$  and  $\alpha 1$  genes. The  $\beta$ -globin gene cluster is located on 11p15.5 and includes five functional genes, 5'- $\epsilon$ - $\gamma$ - $\delta$ - $\beta$ -3' [3].

### 18.2.2 Characteristics of the $\alpha$ -globin and $\beta$ -globin gene loci

Both contain three exons separated by two introns or intervening sequences (IVSs). The  $\alpha$ -globin genes are constitutively expressed owing to their open chromatin configuration with a high G + C content and multiple CpG islands [4]. The exons 1 and 3 of  $\beta$ -globin gene code for non-heme-binding regions, whereas exon 2 codes for heme-binding and  $\beta$ - $\beta$  dimer amino acids. The 5' promoter region, junctions of the exon–intron, and mRNA



**FIGURE 18.1** Organization of the human globin gene clusters on chromosomes 11 ( $\alpha$ -globin gene cluster) and 16 ( $\beta$ -globin gene cluster).

sequence in the 3' untranslated region (3' UTR) contain conserved sequences that are important for  $\beta$ -globin gene expression.

### 18.2.3 Globin gene switch during the fetal to adult transition

Fetal erythrocytes contain predominantly HbF ( $\alpha_2\gamma_2$ ) that extracts oxygen more efficiently from maternal blood. “Hemoglobin switch” is a transition of embryonic–fetal–adult Hb and is mediated through sequential coordinated activation and silencing of embryonic genes. In adults, HbA ( $\alpha_2\beta_2$ ) is  $\sim 97\%$ , HbA<sub>2</sub> ( $\alpha_2\delta_2$ ) is 2%–3.5%, and HbF is  $<1\%$ . HbF levels may vary between normal individuals, which are largely genetically determined [5].

### 18.2.4 $\beta$ -Globin gene expression and its control

The  $\epsilon$ - and  $\gamma$ -globin genes are silenced at early developmental stages, and upregulation of  $\beta$ -globin gene expression relies on no competition for the locus control region (LCR) sequences by  $\gamma$  gene. Downregulation of  $\beta$  gene is noted when the  $\gamma$  gene is upregulated by promoter mutations in cases with nondeletional hereditary persistence of fetal Hb [6]. Occasionally,  $\beta$ -promoter mutations remove competition for the LCR and thereby increase  $\gamma$ - and  $\delta$ -gene expression [7].

## 18.3 The classification and genetics of the thalassemias

The thalassemias are congenital anemias characterized by reduced synthesis rates of one or more of globin subunits of Hb. Cooley and Lee in 1925 [8] described thalassemia as a clinical entity. They are classified according to reduction of a particular globin chain and include  $\alpha$ ,  $\beta$ ,  $\delta\beta$ ,  $\gamma\delta\beta$ ,  $\gamma$ , and  $\delta$  thalassemias. Due to their severity and high frequencies, the  $\beta$  followed by  $\alpha$  thalassemias pose a major public health burden worldwide [9].

### 18.3.1 $\beta$ Thalassemia

Results from either complete ( $\beta^0$ ) or partial absence ( $\beta^+$ ) of  $\beta$ -globin chains. Inheritance is autosomal recessive, that is, both parents of an affected individual with thalassemia major (TM) are  $\beta$ -thalassemia traits ( $\beta$ TTs) who are carriers. Diverse clinical manifestations range from transfusion-dependent state of TM to asymptomatic  $\beta$ TT and intermediate severity cases of thalassemia intermedia (TI). Patients with TM present within 2–3 years of life and require regular red cell transfusions for survival, whereas in TI red cell transfusions are required at reduced frequency.

Clinical features of  $\beta$ TM include failure to thrive, progressive pallor, and on examination they have severe pallor, mild jaundice, and hepatosplenomegaly. If transfusion is not initiated, these children succumb to early high-output heart failure. Patients with TM and TI can show growth retardation, pallor, jaundice, poorly developed musculature, hepatosplenomegaly, leg ulcers, and genu valgum due to inadequate management. Extramedullary hematopoiesis manifests with overt facial changes and tumor-like masses resulting from bone marrow expansion.

Chronic hypertransfusion therapy leads to iron overload and deposition of iron in endocrine organs, liver and heart [10]. Iron chelation therapy is initiated after 15–20 packed red blood cells transfusions.

### 18.3.2 Molecular pathogenesis of $\beta$ thalassemia

Mutations that completely inactivate the  $\beta$ -globin gene cause  $\beta^0$  thalassemia. Primarily, mutations causing reduced production of  $\beta$  globin are classified as  $\beta^+$  or  $\beta^{++}$  thalassemia. This leads to accumulation of excess free  $\alpha$ -globin chains that are the main culprits in disease pathophysiology. Clinical amelioration by coinheritorship of  $\alpha$  thalassemia leads to TI by decreasing free  $\alpha$ -globin chains. Also, coinheritorship of conditions leading to increased  $\gamma$ -globin chain production also results in milder phenotypes by increasing HbF production. Almost 350  $\beta$ -thalassemia mutations, predominantly point mutations within the  $\beta$  gene or its immediate flanking regions [11] have been characterized that are available in an open source database at <http://globin.cse.psu.edu>.

### 18.3.3 Molecular basis of nondeletional $\beta$ thalassemia

#### 18.3.3.1 Mutations that alter gene transcription, that is, mRNA synthesis

These include point mutations occurring within the conserved  $\beta$ -globin promoter (including the CACCC, CCAAT, and ATAA boxes) as well as a 50-nucleotide long stretch in the 5' UTR. These can lead to  $\beta^{++}$  thalassemia alleles that are “silent” with mRNA output of 10%–25% of normal red cell indices [12]. The  $-88\text{ C} > \text{T}$  (HBB:c.  $-138\text{C} > \text{T}$ ) in 20% Africans ([12,13]) and Asian Indians in a single caste [14] have been identified on different haplotypes. Normal RBC indices as well as HbA<sub>2</sub> levels have been noted in about 50% cases of CAP  $+1\text{A} > \text{C}$  (HBB:c.  $-50\text{A} > \text{C}$ ) allele [15,16]. The  $+33\text{C} \rightarrow \text{G}$  (HBB:c.  $-18\text{C} > \text{G}$ ) results in a 33% reduction of  $\beta$  mRNA [17].

Unexplained ethnic variations in phenotypes have been noted with  $-29\text{A} \rightarrow \text{G}$  (HBB:c.  $-79\text{A} > \text{G}$ ) mutation black homozygotes having mild disease [18], whereas Chinese homozygotes have TM [19]. The differences could be due to the  $-158\text{ }^{\text{G}}_{\gamma}\text{C} > \text{T}$  Xmn1 polymorphism, which is known to increased HbF production in the Blacks but is absent in the Chinese chromosome.

#### 18.3.3.2 Mutants that affect mRNA processing

**Mutations in invariant dinucleotides** GT or AG at exon–intron splice junctions can completely eradicate normal splicing, causing  $\beta^0$  thalassemia.

**Mutations in consensus sequences**, which flank the invariant dinucleotides, encompass the last three exonic nucleotides and the first six intronic nucleotides for the 5' donor site; and the last 10 intronic nucleotides and the first exonic nucleotide for the 3' acceptor site. They variably reduce the normal splicing efficiency resulting in phenotypically severe-to-mild  $\beta$  thalassemia. IVS1,  $5\text{G} > \text{C}$  (HBB:c.92 + 5) is commonly found in Indians.

**“Cryptic” splice sites in exons and introns** are sequences analogous to the consensus sequence at splice sites that are not used during normal mRNA processing. Mutations in these sites can create a “similar-to-normal splice site” sequence resulting in aberrant splicing. For example, the IVS1-110  $\text{G} > \text{A}$  (HBB:c.93–21 $\text{G} > \text{A}$ ) mutation, common in the Mediterranean, creates an alternative acceptor AG that lies 19 bp proximal or 5' to the normal acceptor AG of IVS1 [20] resulting in severe  $\beta^+$  thalassemia.

**New donor sites** can be created by substitutions within introns causing  $\beta$  thalassemia such as the IVS 2 position  $654\text{C} > \text{T}$  and  $705\text{T} > \text{G}$  [21].

#### 18.3.3.3 Mutations resulting in abnormal posttranscriptional modification of mRNA

The precursor globin mRNA molecule requires modifications at both ends to attain functionality. A methylated cap structure (m<sup>7</sup>G) is appended to the 5' end while a poly-A tail is added at the 3' end, which is guided by the consensus AATAAA sequence located approximately 20 nucleotides upstream of the poly-A tail that also controls proper cleavage of the primary RNA transcript. Mutations within this AATAAA sequence markedly reduce the cleavage–polyadenylation process efficiency [22] leading to moderately severe  $\beta^+$  thalassemia.

### 18.3.4 Mutants that affect $\beta$ -globin mRNA translation

**Mutations that abolish the initiation codon** (ATG) lead to  $\beta^0$  thalassemia. They could be single base substitutions of any of the three nucleotides A, T, or G [23,24] or may involve insertions within the codon leading to nonsense-mediated decay.



**Introduction of premature stop/termination** codons is a common event that occurs due to mutations directly changing a coding codon to a stop codon or a frameshift mutation causing misreading of the genetic code by insertion or deletion of nucleotide/s [25]. Examples include codon 39 (CAG to TAG  $\beta$  39(C5) Gln > Stop HBB: c.118C > T) mutation that is common in the Mediterranean and in Sardinia [26].

## 18.4 Gene deletions in $\beta$ thalassemia

Major gene deletions in  $\beta$  thalassemias are rare and are classified based on whether they involve the  $\beta$ -globin gene alone or also include genes/regions lying upstream of the gene.

### 18.4.1 Deletions restricted to the $\beta$ -globin gene

Deletions that involve only the  $\beta$ -globin gene usually range from approximately 105 bp to  $\sim 67$  kb in their size and are  $\beta^0$  thalassemia [24]. The 619 bp deletion involving the 3' end of the  $\beta$ -globin gene is common among Indians, seen in nearly 30% cases [27]. Other deletions vary widely in extents but share in common the loss of a  $\beta$  promoter region from positions  $-125$  to  $+78$  (numbers relative to the mRNA cap site). Heterozygotes in all of these display high HbA<sub>2</sub> levels with variable elevations of HbF. Deletion of  $\beta$  promoter locus possibly removes competition for upstream  $\beta$ -LCR and allows greater interaction of LCR with transcription factors, leading to enhancement of expressions of the  $\delta$  and  $\gamma$  genes [26].

### 18.4.2 Upstream deletions and $(\epsilon\gamma\delta\beta)^0$ thalassemia

These affect the  $\beta$ -LCR and downregulate expression of  $\beta$ -globin gene along with all its linked genes in the chromosome 11p cluster, resulting in  $(\epsilon\gamma\delta\beta)^0$  thalassemia. Deletions can remove all or nearly all of the cluster or can remove the upstream LCR, but the HBB gene is itself left intact [24]. Recognition of private rare upstream deletions has enhanced our understanding of the importance of long-range regulatory elements in  $\beta$ -globin locus control [28]. Degree of anemia and hemolysis are variable even for identical mutations within a single family [9]. Since only heterozygotes have been identified, this suggests that homozygotes die during early gestation.

## 18.5 Other less common, specific molecular causes of $\beta$ thalassemia

### 18.5.1 Dominant $\beta$ thalassemia

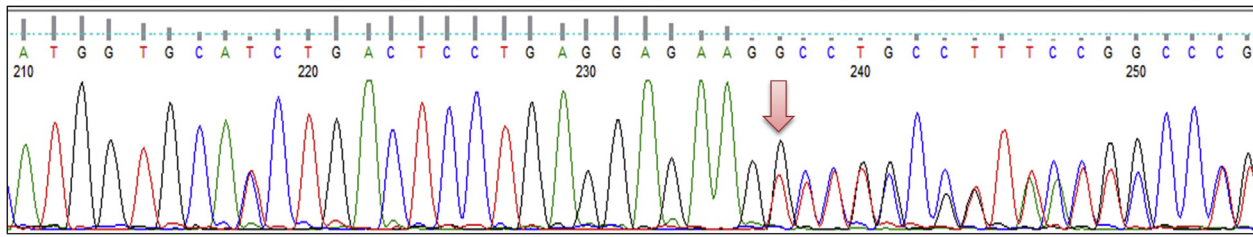
The occurrence is autosomal dominant/sporadic and cases show raised HbA<sub>2</sub>, severe marrow dyserythropoiesis with inclusion bodies in erythroblasts (composed of  $\beta$  and  $\alpha$  chains). Hence, this condition is also known as "inclusion body  $\beta$  thalassemia" [9]. The molecular defects are variable and lead to the synthesis of unstable  $\beta$ -globin chains, which undergo rapid degradation. These unstable  $\beta$  chains precipitate, and excess  $\alpha$  chains cause ineffective erythropoiesis leading to severe phenotypes. Gene level defects include indel mutations (e.g., Hb Korea and Hb Gunma), missense events [e.g., Hb Terre Haute ( $\beta$ 106 Leu  $\rightarrow$  Arg)], nonsense events (GAA  $\rightarrow$  TAA codon 121) and frameshifts causing aberrant splicing (especially those ending late in exon 3) [26].

### 18.5.2 Silent and almost silent $\beta$ -thalassemia trait

Silent  $\beta$ TT such as the CAP + 1A > C has been described. Other silent mutations include the Greek + 1480C > G mutation, the  $-92$ C > T (HBB:c.  $-142$ C > T) and  $-101$ C > T (HBB:c.  $-151$ C > T),  $\sim 50\%$  of the heterozygotes are silent. Acquired abnormalities such as zidovudine therapy [29] or liver disease can elevate MCV in some patients, while in others a mild decline in HbA<sub>2</sub> may result from very severe iron deficiency [30].

### 18.5.3 Trans acting mutations associated with $\beta$ thalassemia

This rare pathology remains obscure despite complete sequencing of the  $\beta$ -globin gene and its flanking regions. *Trans* acting mutations have been detected in some cases such as *XPD* (trichothiodystrophy with a  $\beta$ TT phenotype [31]), *GATA-1* (X-linked  $\beta$  thalassemia and thrombocytopenia [32]), and *KLF1*.



**FIGURE 18.2** Sequencing chromatogram for heterozygous Frameshift 8/9 + G (HBB:c.27\_28insG) depicted by the arrow.

#### 18.5.4 Uniparental isodisomy/somatic deletion of $\beta$ -globin gene

Mosaic somatic deletions in *HBB* gene (i.e., in a subpopulation of erythroblasts) leading to moderately severe anemia and hepatosplenomegaly were described [33] in patients who also are germline carriers of another typical  $\beta$ -thalassemia mutation. Uniparental isodisomy in chr.11p has also been described [34].

### 18.6 Laboratory diagnosis of $\beta$ thalassemia

Screening tests to identify  $\beta$ TT include Hb estimation, red cell indices, and quantitation of HbF and HbA<sub>2</sub> using cation-exchange Hb high-performance liquid chromatography (HPLC), isoelectric focusing, or capillary electrophoresis [35]. Increased HbA<sub>2</sub> is the diagnostic hallmark of  $\beta$ TT where the level is usually between 4% and 7%. Borderline levels between 3.6% and 3.9% need to be interpreted with caution and a DNA analysis should be advised specially if the partner is a  $\beta$ TT [36].  $\beta$ -Thalassemia mutations are analyzed using a PCR-based amplification refractory mutation system (ARMS), reverse dot-blot, or DNA sequencing (Fig. 18.2) using genomic DNA from blood leukocytes.

### 18.7 Prevention of $\beta$ thalassemia

Preventive strategies for  $\beta$  thalassemia include population screening for carriers at either premarital or preconception stage followed by genetic counseling and fetal testing.

#### 18.7.1 Screening for $\beta$ -thalassemia trait

Voluntary screening programs and counseling should be offered primarily by creating awareness of genetic risks via mass media. Screening can be carried out at primary health-care level, schools, or young adults before conception, specifically in prenatal women. These strategies have been very successful in countries such as Sardinia, Cyprus, and Greece.

#### 18.7.2 Prenatal diagnosis

Since 1974, when the first prenatal diagnosis was carried out, this has become a major approach to prevent birth of homozygous  $\beta$  thalassemia. Strategy has evolved from obtaining fetal blood for the analysis of globin chain synthesis to specific mutation detection. Genetic counseling is required based on ethical, legal, social, and financial considerations [37]. The fetal sample of choice is the early chorionic villous sampling taken at 10–12 weeks of gestation. Amniocentesis is also performed if the couple reports late. Many centers in the world, including India, have adopted this method of prevention of  $\beta$  thalassemia [14,38,39].

#### 18.7.3 Preimplantation genetic diagnosis

Preimplantation genetic diagnosis (PGD) is an approach where the congenital disorder is avoided through the selection of unaffected gametes or embryos before pregnancy [40]. This entails an in vitro process and offers the advantage that an abortion is avoided. For performing PGD the sample analyzed is a polar body or a single cell

taken from 4 to 8 cell blastomere. DNA is extracted from the single cell and PCR performed for the mutations in the case of  $\beta$  thalassemia. This approach has become popular in the West and though there are centers in south-east Asia, the costs become the prohibitive factor to be available for routine use.

#### 18.7.4 Noninvasive prenatal diagnosis by analyzing cell-free circulating fetal DNA in the maternal blood

Isolating fetal cells from maternal circulation was introduced in 1996, and many studies were described. However, the yield of fetal cells was very low, and the technique was cumbersome [41–43]. Cell-free fetal DNA from maternal plasma is currently being technically refined for obtaining DNA as a noninvasive method [44,45]. Currently, research is underway to use noninvasive prenatal test for  $\beta$  thalassemia utilizing a novel DNA probe capture assay by next-generation sequencing.

Recently “microfluidics digital PCR” has found a higher than expected fraction of fetal DNA in maternal plasma, and this is helpful in quantitating the percentage of fetal DNA fraction [46].

### 18.8 $\alpha$ Thalassemia

$\alpha$  thalassemia is the commonest asymptomatic single-gene disorder in the world [9]. Deletions occur commonly compared to point mutations. Many *Alu*-family repeats are present throughout the  $\alpha$ -globin gene cluster loci, and these sites are frequently associated with DNA recombination causing variable sized deletions. Rarely, deletions of the upstream regulatory elements, which have a major role in controlling  $\alpha$ -globin gene expression, can be encountered.

#### 18.8.1 Classification and clinical phenotypes of $\alpha$ thalassemia

Normal  $\alpha$  genes are duplicated and genotype is represented as  $\alpha\alpha/\alpha\alpha$  [47]. Four types of  $\alpha$  thalassemia are found that range from mild to severe clinical phenotypes, which include (1)  $\alpha^-/\alpha\alpha$  (silent carrier/ mild  $\alpha$ -thalassemia/ $\alpha^+$  thalassemia trait) due to commonly 3.7 or uncommonly 4.2 kb deletion in one allele; (2) homozygous  $\alpha^+$  thalassemia ( $\alpha^-/\alpha^-$ )/  $\alpha^0$  trait ( $-/\alpha\alpha$ ); (3) HbH disease ( $\alpha^-/-$ ) [48]; and (4) hydrops fetalis or  $\alpha$  TM/ Barts Hb ( $-/-$ ), which is usually incompatible with life, unless fetal blood transfusions are initiated.

A mutation in the  $\alpha$  gene is designated as “T,” which inactivates one of the pairs is an uncommon phenomenon compared to deletions. Nondeletional homozygous state is designated as  $\alpha^T\alpha/\alpha^T\alpha$  where patients are more symptomatic with chronic hemolytic anemia [49] and heterozygous state as  $\alpha^T\alpha/\alpha\alpha$  is asymptomatic [50]. The overwhelming majority affect the  $\alpha 2$  gene and a few are described in *cis* forms with deletions ( $-\alpha^T$ ) [47].

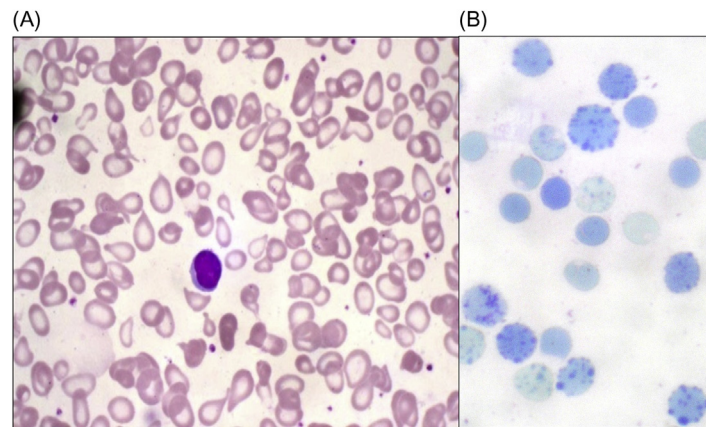
#### 18.8.2 Molecular pathology of deletional and nondeletional $\alpha$ thalassemias

The  $\alpha 2$ - and  $\alpha 1$ -globin genes are embedded within two 4 kb homologous units, which have homologous sub-segments X, Y, and Z. The Z segments are 3.7 kb away, and reciprocal recombination results in only one  $\alpha$  gene (rightward deletion— $\alpha^{3.7}$ ), and the other allele has three  $\alpha$  genes ( $\alpha\alpha\alpha^{3.7}$ ). Recombination between the X boxes that are 4.2 kb apart results in leftward deletion— $\alpha^{4.2}$  and an  $\alpha\alpha\alpha^{4.2}$  allele. The  $\alpha^+$  deletional forms (including  $-\text{MED}$ ) are in high frequencies throughout the tropical belt and are in highest frequencies in endogamous communities. The  $\alpha^0$  deletions are highest in southeast Asia ( $-\text{SEA}$ ,  $-\text{THAI}$ ,  $-\text{FIL}$ ) [51]. In north Indians,  $\alpha^+$  thalassemia was first documented in patients of TI [52]. A detailed study carried out showed that amongst north Indians, 12%–13% of the population have  $\alpha^+$  thalassemia and triplicated  $\alpha$  gene is 3%. The vast majority comprising 98% are— $\alpha^{3.7}$  and only 2% are— $\alpha^{4.2}$  allele [53,54]. A study from South India showed that  $-\text{SA}$  was common in Indians [55]. They often function as phenotype modifiers in other hemolytic anemias [56].

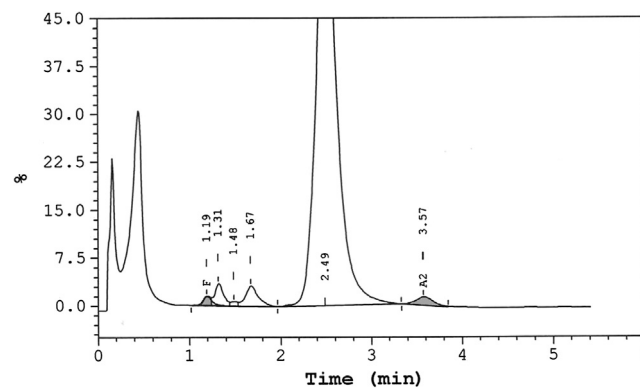
Nondeletional Hb Koya Dora  $\alpha 142$ , Term  $\rightarrow$  Ser ( $\text{TAA} > \text{TCA}$  in  $\alpha 2$ ) has been found in central India [57]. Hb Sallanches [ $\alpha 104(\text{G11})\text{Cys} \rightarrow \text{Tyr}$  ( $\text{TGC} > \text{TAC}$ ) ( $\alpha 2$ )] and Hb Sun Prairie [ $\alpha 130(\text{H13})\text{Ala} \rightarrow \text{Pro}$ ,  $\text{GCT} > \text{CCT}$  ( $\alpha 2$ )] in Asian Indians [58]. Hb Constant Spring  $\alpha 142$ , Term  $\rightarrow$  Gln ( $\text{TAA} > \text{CAA}$  in  $\alpha 2$ ) is found in frequencies of 5%–8% in southeast Asia.

#### 18.8.3 Laboratory diagnosis of $\alpha$ -deletions, point mutations and triplications

The suspicion of symptomatic  $\alpha$  disease should be considered after excluding common conditions such as iron deficiency anemia and  $\beta$  thalassemia syndromes. Investigations include hemogram with red cell indices which



**FIGURE 18.3** (A) Peripheral blood in a patient with homozygous Hb Sallanches [ $\alpha$  2 104(G11) Cys>Tyr HBA2:c.314G>A] showing marked hypochromic microcytosis and anisopoikilocytosis. (B) “Golf-ball” type HbH inclusions in red cells stained supravitaly with brilliant cresyl blue.



**FIGURE 18.4** Cation-exchange high-performance liquid chromatography (CE-HPLC) pattern of a case of HbH disease showing twin preintegration peaks and reduced HbA<sub>2</sub>.

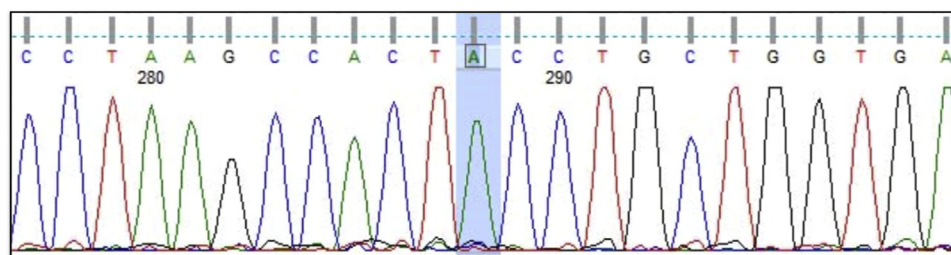
shows anisopoikilocytosis (Fig. 18.3A), screening for HbH inclusions (Fig. 18.3B) and HPLC where twin preintegration peaks with low HbA<sub>2</sub> levels are seen (Fig. 18.4). Hemoglobin electrophoresis at an alkaline pH shows a fast-moving band of Hb Barts. Tests for unstable Hbs include isopropanol solubility test and heat instability test which are positive with HbH disease.

A multiplex gap-PCR on genomic DNA includes testing for  $-\alpha^{3.7}$ ,  $-\alpha^{4.2}$ ,  $_{-SEA}$ ,  $_{-MED}$ ,  $_{-SA}$ ,  $\alpha\alpha^{THAI}$ ,  $\alpha\alpha^{FIL}$  and  $-(\alpha)^{20.5}$  double-gene deletions [51]. Presence of  $\alpha^0$  allele is suspected whenever the  $\alpha_2$  band is missing. One of the parents of a case with HbH disease shows a normal pattern by gap-PCR. For triplications PCR across the junction of the crossover is carried out in the same reaction  $\alpha\alpha\alpha^{3.7}$  and  $\alpha\alpha\alpha^{4.2}$  is carried out [59].  $\alpha$ -Globin gene sequencing is done to identify point mutations and PCR is carried out as a two stage PCR similar to a nested PCR (Fig. 18.5). For identifying variable sized deletions which are novel, multiplex ligation-dependent probe amplification (MLPA) analysis needs to be carried out (Fig. 18.6) using  $\alpha$ -Globin<sup>XS</sup> MLPA kit from ServiceXS, Leiden, Netherlands or MRC, Holland.

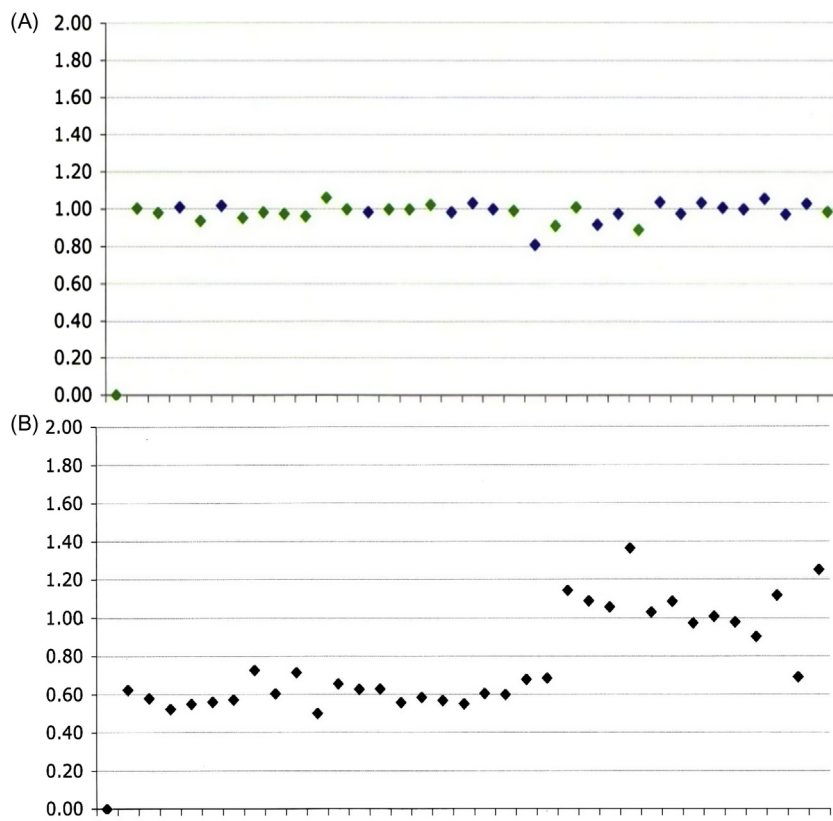
#### 18.8.4 *Thalassemia intermedia: Molecular genetics and genotype–phenotype correlation*

The clinical features of TI are variable mild-to-severe phenotypes, though milder than TM. The mild clinical characteristics occur primarily due to three different mechanisms [60,61] which are as follows:

1. Inheritance of a mild ( $\beta^+$ ) or silent  $\beta$ -chain ( $\beta^{++}$ ) mutation with some  $\beta$  output
2. Coinheritance of factors associated with enhanced  $\gamma$ -globin chain production, for example,  $Xmn1^G\gamma$  polymorphism; *Trans* acting quantitative trait loci for HbF on Xp22.2-p22.3, 6q23, 8q, and 2p15
3. Coinheritance of  $\alpha$  thalassemia leading to reduction of  $\alpha$ : $\beta$ -chain imbalance.



**FIGURE 18.5**  $\alpha 2$  Sequencing shows homozygosity for Hb Sallanches ( $\alpha 2$  codon 104G > A; Cys  $\rightarrow$  Tyr; G11 Helix). *Hb*, Hemoglobin.



**FIGURE 18.6** MLPA pattern of the three new deletions noted in a patient from north India. (A) Normal pattern, (B) probes 1–22 deleted. *MLPA*, Multiplex ligation-dependent probe amplification.

Occasionally, heterozygous  $\beta$  thalassemia with coinheritance of additional  $\alpha$ -globin genes ( $\alpha\alpha\alpha/\alpha\alpha$ ,  $\alpha\alpha\alpha/\alpha\alpha\alpha$ ,  $\alpha\alpha\alpha\alpha/\alpha\alpha$ ,  $\alpha\alpha\alpha\alpha/\alpha\alpha\alpha\alpha$ ) can exacerbate the phenotype by causing more severe globin chain imbalance [62]. Rarely, dominantly inherited  $\beta$  thalassemia or *KLF1* mutations can cause TI [63,64].

## 18.9 Qualitative defects (structural hemoglobin variants and other abnormalities)

These include the clinically and epidemiologically important Hbs such as HbS, HbE, and HbC as well as clinically innocuous variants that are primarily of laboratory importance, such as HbQ-India and HbJ-Meerut [65,66]. Only the former ones are discussed here.

### 18.9.1 Sickle-cell hemoglobin

Sickle-cell Hb affects millions of people across the world, and an estimated 400,000 homozygous persons are born annually [67]. This monogenic disorder is caused by an A > T single-nucleotide substitution in the  $\beta$ -globin



gene that produces the abnormal HbS. Deoxygenated HbS tends to polymerize, resulting in morphologically observable sickling of red cells. Sickled erythrocytes are dehydrated and display increased endothelial adhesion leading to small vessel thrombosis and hemolysis. These result clinically in painful crises and features of organ damage due to hypoxia [68].

There is significant heterogeneity among patients with sickle-cell disease, accounted for by the various combinations it can present as (homozygous SS, double heterozygous sickle- $\beta$  thalassemia) as well as by  $\alpha$  thalassemia, regulators of HbF, genes involved in vascular tone and repair regulation and environmental and infectious factors [68]. Diagnosis of a sickling syndrome relies upon demonstration of the abnormal Hb by alkaline pH electrophoresis or HPLC and confirmation by a second test such as the sickle solubility test or slide-based test for sickling using sodium metabisulfite or dithionite. Parental studies may be required for distinction between genotypes [69].

Management of sickle crisis includes hydration and supportive therapy including blood transfusions and pain management. Hydroxyurea also results in clinical improvement attributable to increased HbF and reduction in dense undeformable sickle cells, highly adhesive sickle-cell reticulocytes and granulocytes, all of which alter disease severity independent of sickling phenomena [68]. Stem-cell transplant has been explored for very severe cases, and recently the US-FDA has licensed L-glutamine, a nicotinamide adenine dinucleotide (NAD) precursor that alleviates red cell oxidative stress, for reducing painful crises in these patients [70].

### 18.9.2 Hemoglobin E

HbE is a  $\beta$  chain hemoglobinopathy that is caused by a single base substitution at position 26 of the  $\beta$  chain (GAG > AAG, glutamic acid to lysine). It is a mildly unstable Hb that is common in southern and southeastern Asian countries. The instability results from activation of a cryptic splice site that reduces the level of normally spliced  $\beta^E$  chain. Heterozygotes are mostly asymptomatic or mildly anemic with microcytic hypochromic red cell indices. The variant, which co-elutes with HbA2 on HPLC comprises 25%–30% of total Hb and the percentage reduces with coinherited  $\alpha$  thalassemia and iron deficiency. Homozygotes have 80%–95% of the variant Hb yet resemble  $\beta$ TT clinically with normal to mildly reduced Hb and prominent microcytosis with target cells on a smear [71].

The major clinically important states are when HbE combines with  $\beta$  thalassemia (E- $\beta$  thalassemia) with or without coexisting  $\alpha$  thalassemia. This is a disease of great clinical heterogeneity, ranging from very mild TI to an almost transfusion-dependent state. Clinical features include symptomatic anemia, jaundice, hepatomegaly, splenomegaly, growth retardation, hemolytic facies, and leg ulcers. Potential complications include iron overload, extramedullary hematopoietic masses, and increased risks of thrombosis, infections, and diabetes mellitus.

Coinheritance of heterozygous and homozygous HbE with HbH disease give rise to Hb AE Bart's and Hb EF Bart's disease respectively, both of which are associated with a TI phenotype of variable severity [71].

### 18.9.3 Hemoglobin C

HbC is a  $\beta$  globin variant with highest prevalence in western Africa and people of African descent in North America and Europe, especially the southern parts. Its major pathogenic effects result from its ability to crystallize in the oxyhemoglobin state, with resolubilization on deoxygenation. Heterozygotes are typically asymptomatic while homozygotes show a mild, chronic hemolytic anemia, often with splenomegaly, and sometimes gallstones. The blood smear displays marked microcytosis, targetemia and increase irregularly contracted cells. HbC crystals are 6–10  $\mu$ m long and 2–3  $\mu$ m wide hexagonal or tetragonal inclusions usually seen in dehydrated cells that are devoid of cytoplasm and are more common postsplenectomy. Coinheritance of HbC with  $\beta$  thalassemia leads to a moderate to severe TI phenotype [9].

### 18.9.4 Hemoglobin M or methemoglobinemic hemoglobin variants

Inherited methemoglobinemia can be due to oxidant damage (especially in G6PD-deficient persons) or inherited deficiency of NADH (reduced NAD)-cytochrome b5 reductase. It can also occur due to an abnormal variant Hb that shows an increased tendency to oxidize to methemoglobin and is therefore called HbM. HbMs include  $\beta$ -globin variants such as Hb M-Saskatoon and Hb M-Milwaukee-1,  $\alpha$ -globin variants such as Hb M-Boston and Hb Auckland and also  $G\gamma$ -globin variants such as Hb F-M-Osaka and Hb F-M-Fort

Ripley [72]. HbMs typically arise from mutations that alter the amino acids of the heme-binding pocket of the Hb molecule. Most such mutations lead to tyrosine replacing a proximal or distal histidine residue [9]. Amino acid substitutions at these critical heme-binding sites result in prolonged and irreversible oxidation of ferrous ( $\text{Fe}^{2+}$ ) iron to the ferric ( $\text{Fe}^{3+}$ ) state in the variant Hbs and this is sufficient to alter several critical properties including oxygen affinity (P50), electrophoretic mobility, chromatographic retention time and spectrophotometric absorption spectra [73].

Clinical features include congenital cyanosis (that may go unnoticed in darker skinned individuals) and abnormalities on blood gas analysis. Management is usually conservative for these mostly cosmetic abnormalities [73]. Some variants, such as HbM-Hyde Park may be unstable as well.

### 18.9.5 Unstable hemoglobins

Hemoglobin molecules may become unstable at a molecular level due to abnormalities that affect the hydrophobic heme pocket, interfere with the  $\alpha$  helical or tertiary structures, or impact the interactions of the  $\alpha$  and  $\beta$  subunits. Clinically, they are characterized by dominant inheritance, at least mild anemia and the laboratory features of reticulocytosis, Heinz body formation, and positive heat and isopropanolol instability tests. Examples of common unstable Hbs include Hb Zurich and Hb Koln. About one-thirds of the unstable Hbs may display high oxygen affinity as well, but, since they also cause hemolysis, anemia rather than polycythemia is seen [74,9].

### 18.9.6 High-oxygen affinity hemoglobins

These mostly dominantly inherited variants present as polycythemia at a young age and an increased risk of thrombosis in later life. The underlying molecular abnormality results in either reduced affinity for 2,3-BPG or any impairment of the subunits that causes abnormal stabilization of the oxy- or taut configuration. In the past, mislabeling of high-oxygen affinity Hbs as polycythemia vera with subsequent administration of inappropriate therapies was common. Diagnosis requires both Hb HPLC and electrophoresis for a variant peak/band as well as estimation of P50 and the oxygen dissociation curve. Tests for instability may help as some of these variants are unstable. Examples of such high-oxygen affinity Hbs include Hb Chandigarh, Hb Rainier, Hb Andrew-Minneapolis and Hb McKees Rocks [9,75].

### 18.9.7 Low-oxygen affinity hemoglobins

A low oxygen affinity results in better oxygen delivery to the tissues even at low partial pressures, hence meaning that anemia is well tolerated. Even homozygous sickle-cell patients display this phenomenon, resulting in lessening of the erythropoietic drive. Other examples include Hb Kansas and Hb Beth Israel, which show mild anemia without major symptoms and have reduced oxygen saturation on pulse oximetry [9].

### 18.9.8 Defects of erythroid heme biosynthesis

Heme biosynthesis is a tightly regulated pathway that involves eight enzymes located in the cytosol or the mitochondria. The pathway begins with  $\delta$ -aminolevulinic acid synthase-2 (ALAS-2) and ends with ferrochelatase, and defects in both the enzymes can cause diseases with significant red cell lineage involvement.

X-linked sideroblastic anemia is characterized by the presence of ring sideroblasts in the bone marrow due to the presence of iron-laden mitochondria in these cells. Patients present mostly in childhood with anemia and variable degrees of hepatosplenomegaly. The genetic defects are heterogeneous, mostly loss-of-function missense mutations in *ALAS-2*, but also other genes such as *ABCB7* and *GLRX5* can be involved.

A partial ferrochelatase deficiency causes erythropoietic porphyria and results in elevated unbound protoporphyrin in erythrocytes with secondary accumulation in liver and skin. Clinical features include recessive inheritance, photosensitivity, mild microcytic anemia, and, in a minority of patients, liver disease [76].

## 18.10 Summary

1. Clinical and laboratory research on Hb has formed the basis of much of our current knowledge of molecular medicine.
2. Disorders of Hb include globin chain defects resulting in the quantitative thalassemia syndromes as well as structurally and functionally variant Hbs that may be unstable, cause sickling, methemoglobinemia, or have high or low oxygen affinity.
3. The thalassemias represent the commonest monogenic disorders worldwide and occur in high frequencies in the malaria belt. The severe forms are characterized by anemia requiring transfusional support, and complications thereof including growth failure, iron overload, heart failure, infections, bone disease, etc.
4.  $\beta$ TT screening can be done by detecting high HbA<sub>2</sub> using HPLC, microcolumn chromatography or capillary electrophoresis.  $\beta$ -Globin (*HBB*) mutations can be detected by reverse dot-blot analysis, ARMS-PCR or Sanger sequencing.
5.  $\alpha$ -globin (*HBA1* and *HBA2*) genes' defects are most accurately identified by molecular tests, an exception being HbH disease that yields a "golf-ball" appearance on supravital staining, preintegration peaks on HPLC and a fast-moving band on alkaline pH electrophoresis.
6. Multiplex gap-PCR or multiplex ligation-dependent probe amplification identifies common  $\alpha$ -globin gene deletions while nondeletional  $\alpha$  thalassemia is detected by specific  $\alpha 1$ - or  $\alpha 2$ -globin gene sequencing.
7. The structural variants include HbS (sickle Hb), HbE, and HbC among others and pose region and ethnic group-specific problems.
8. Homozygous or compound heterozygous sickle-cell states lead to anemia, vascular occlusions, painful crises, organ infarctions, chest infections, and reduced longevity.
9. Most of the thalassemias and hemoglobinopathies are inherited as autosomal recessive disorders and most variants leading to unstable Hb, methemoglobinemia or high- or low-oxygen affinity Hbs present as either sporadic cases or autosomal dominant inheritance.

## References

- [1] Perutz MF. Structure and mechanism of hemoglobin. *Br Med Bull* 1976;32:195–208.
- [2] Higgs DR, Thein SL, Wood WG. Human haemoglobin. In: Weatherall DJ, Clegg B, editors. *The thalassaemia syndromes*. 4th ed. Oxford: Blackwell Science; 2001.
- [3] Thein SL. Genetic insights into the clinical diversity of beta thalassaemia. *Br J Haematol* 2004;124:264–74.
- [4] Craddock C.F, Vyas P, Sharpe JA, Ayyub H, Wood WG, Higgs DR. Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. *EMBO J* 1995;14:1718–26.
- [5] Garner C, Tatu T, Reittie JE, Littlewood T, Darley J, Cervino S, et al. Genetic influences on F cells and other hematological variables: a twin heritability study. *Blood* 2000;95:342–6.
- [6] Wood WG. Hereditary persistence of fetal hemoglobin and delta beta thalassemia. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. *Disorders of hemoglobin: genetics, pathophysiology, and clinical management*. Cambridge University Press; 2001. p. 356–88.
- [7] Thein SL, Craig JE. Genetics of Hb F/F cell variance in adults and heterocellular hereditary persistence of fetal hemoglobin. *Hemoglobin* 1998;22:401–14.
- [8] Cooley TB, Lee P. A series of cases of splenomegaly in children with anemia and peculiar bone changes. *Trans Am Pediatr Soc* 1925;37:29.
- [9] Bain BJ, editor. *Haemoglobinopathy diagnosis*. 2nd ed. Malden, MA: Blackwell Publishing Ltd; 2006.
- [10] Borgna-Pignatti C, Rugolotto S, De Stefano P, et al. Survival and complications in patients with thalassemia major treated with transfusion and deferoxamine. *Haematologica* 2004;89:1187–93.
- [11] Giardine B, Borg J, Higgs DR, Peterson KR, Philipson S, Maglott D, et al. Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nat Genet* 2011;43:295–301.
- [12] Gonzalez-Redondo JM, Stoming TA, Kutlar A, Kutlar F, Lanclos KD, Howard EF, et al. A C→T substitution at nt –101 in a conserved DNA sequence of the promoter region of the  $\beta$ -globin gene is associated with "silent"  $\beta$ -thalassemia. *Blood* 1989;73:1705–11.
- [13] Orkin SH, Antonarakis SE, Kazazian Jr. HH. Base substitution at position -88 in a  $\beta$ -thalassaemia globin gene. *J Biol Chem* 1984;259:8679–81.
- [14] Garewal G, Das R, Ahluwalia J, Marwaha RK, Varma S. Nucleotide -88 (C-T) promoter mutation is a common  $\beta$ -thalassemia mutation in the Jat Sikhs of Punjab, India. *Am J Haematol* 2005;79:252–6.
- [15] Garewal G, Das R, Awasthi A, Ahluwalia J, Varma N, Marwaha RK. The clinical significance of the spectrum of interactions of CAP + 1 (A>C), a silent beta globin gene mutation, with other beta thalassemia mutations and globin gene modifiers in North Indians. *Eur J Haematol* 2007;79:417–21.
- [16] Rangan A, Sharma P, Dadu T, Saxena R, Verma IC, Bhargava M.  $\beta$ -Thalassemia mutations in subjects with borderline HbA<sub>2</sub> values: a pilot study in North India. *Clin Chem Lab Med* 2011;49:2069–72.
- [17] Ho PJ, Rochette J, Fisher CA, Wonke B, Jarvis MK, Yardumian A, et al. Moderate reduction of  $\beta$  globin gene transcript by a novel mutation in the 5' untranslated region: a study of its interaction with other genotypes in two families. *Blood* 1996;87:1170–8.

- [18] Safaya S, Rieder RF, Dowling CE, Kazazian HHJ, Adams JGI. Homozygous  $\beta$ -thalassemia without anemia. *Blood* 1989;73:324–8.
- [19] Huang SZ, Wong C, Antonarakis SE, Ro-Lein T, Lo WHY, Kazazian HHJ. The same TATA box  $\beta$ -thalassemia mutation in Chinese and US Blacks: another example of independent origins of mutation. *Hum Genet* 1986;74:162–4.
- [20] Busslinger M, Moschonas N, Flavell RA.  $\beta^+$  thalassemia: aberrant splicing results from a single point mutation in an intron. *Cell* 1981;27:289–98.
- [21] Orkin SH, Kazazian HHJ, Antonarakis SE, Ostrer H, Goff SC, Sexton JP. Abnormal RNA processing due to the exon mutation of  $\beta$ E-globin gene. *Nature* 1982;300:768–9.
- [22] Orkin SH, Cheng T-C, Antonarakis SE, Kazazian HHJ. Thalassemia due to a mutation in the cleavage-polyadenylation signal of the human  $\beta$ -globin gene. *EMBO J* 1985;4:453–6.
- [23] Forget BG, Steinberg MH, Forget BG, Higgs DR, Nagel RL. Molecular mechanisms of  $\beta$  thalassemia. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. *Disorders of hemoglobin: genetics, pathophysiology, and clinical management*. Cambridge: Cambridge University Press; 2001. p. 252–76.
- [24] Thein SL, Wood WG. The molecular basis of  $\beta$  thalassemia,  $\delta\beta$  thalassemia, and hereditary persistence of fetal hemoglobin. In: Steinberg MH, Forget BG, Higgs DR, Nagel RL, editors. *Disorders of hemoglobin: genetics, pathophysiology, and clinical management*. Cambridge: Cambridge University Press; 2009. p. 323–56.
- [25] Hall GW, Thein SL. Nonsense codon mutations in the terminal exon of the  $\beta$ -globin gene are not associated with a reduction in  $\beta$ -mRNA accumulation: a mechanism for the phenotype of dominant  $\beta$ -thalassemia. *Blood* 1994;83:2031–7.
- [26] Thein SL. The molecular basis of  $\beta$ -thalassemia. *Cold Spring Harbor Perspect Med* 2013;3(5):a011700. Available from: <https://doi.org/10.1101/cshperspect.a011700>.
- [27] Varawalla NY, Old JM, Weatherall DJ. Rare  $\beta$ -thalassaemia mutations in Asian Indians. *Br J Haematol* 1991;79:640–4.
- [28] Kioussis D, Vanin E, de Lange T, Flavell RA, Grosveld FG.  $\beta$ -globin gene inactivation by DNA translocation in  $\gamma\beta$ -thalassaemia. *Nature* 1983;306:662–6.
- [29] Bhagat P, Sachdeva RK, Sharma P, Sachdeva MUS, Chhabra S, Sharma A, et al. Effect of antiretroviral therapy on hemoglobin A<sub>2</sub> values can have implications in antenatal beta thalassemia screening programs. *Infect Dis (Lond)* 2016;48(2):122–6.
- [30] Sharma P, Das R, Trehan A, Bansal D, Chhabra S, Kaur J, et al. Impact of iron deficiency on hemoglobin A<sub>2</sub>% in obligate  $\beta$ -thalassemia heterozygotes. *Int J Lab Hematol* 2015;37:105–11.
- [31] Viprakasit V, Gibbons RJ, Broughton BC, Tolmie JL, Brown D, Lunt P, et al. Mutations in the general transcription factor TFIID result in  $\beta$ -thalassaemia in individuals with trichothiodystrophy. *Hum Mol Genet* 2001;10:2797–802.
- [32] Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH, Raskind WH. X-linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* 2002;100:2040–5.
- [33] Galanello R, Perseu L, Perra C, Maccioni L, Barella S, Longinotti M, et al. Somatic deletion of the normal  $\beta$ -globin gene leading to thalassaemia intermedia in heterozygous  $\beta$ -thalassaemic patients. *Br J Haematol* 2004;127:604–6.
- [34] Chang JG, Tsai WC, Chong IW, Chang CS, Lin CC, Liu TC.  $\beta$ -thalassaemia major evolution from  $\beta$ -thalassaemia minor is associated with paternal uniparental isodisomy of chromosome 11p15. *Haematologica* 2008;93:913–16.
- [35] Sreedharanunni S, Sharma P, Murgai P, Chhabra S, Das R. Automated alkaline-pH electrophoresis followed by densitometry does not correlate with cation-exchange (CE)-HPLC in quantification of HbA<sub>2</sub> and variant hemoglobins. *Clin Chem Lab Med* 2015;53:e301–3.
- [36] Sharma P, Das R. Cation-exchange high-performance liquid chromatography for variant hemoglobins and HbF/A<sub>2</sub>: what must hematopathologists know about methodology? *World J Methodol* 2016;6:20–4.
- [37] Loukoupolis D. Haemoglobinopathies in Greece. *Indian J Med Res* 2011;134:572–6.
- [38] Thakur (Mahadik) C, Vaz F, Banerjee M, Kapadia C, Natrajan PG, Yagnik H, et al. Prenatal diagnosis of  $\beta$ -thalassaemia and other hemoglobinopathies in India. *Prenatal Diagn* 2000;20:194–201.
- [39] Colah RB, Gorakshakar AC, Nadkarni A. Invasive & non-invasive approaches for prenatal diagnosis of haemoglobinopathies: experiences from India. *Indian J Med Res* 2011;134:552–60.
- [40] Hussey ND, Davis T, Hall JR, et al. Preimplantation genetic diagnosis for  $\beta$ -thalassaemia using sequencing of single cell PCR products to detect mutations and polymorphic loci. *Mol Hum Reprod* 2002;8:1136–43.
- [41] Cheung MC, Goldberg JD, Kan YW. Prenatal diagnosis of sickle cell anemia and thalassemia by analysis of fetal cells in maternal blood. *Nat Genet* 1996;14:264–8.
- [42] Steele CD, Wapner RJ, Smith JB, Haynes MK, Jackson LG. Prenatal diagnosis using fetal cells isolated from maternal peripheral blood. *Clin Obstet Gynecol* 1996;39:801–13.
- [43] Takabayashi H, Kuwabara S, Ukita T, Ikawa K, Yamafuji K, Igarashi T. Development of non-invasive fetal DNA diagnosis from maternal blood. *Prenat Diagn* 1995;15:74–7.
- [44] Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, et al. Presence of fetal DNA in maternal plasma and serum. *Lancet* 1997;350:485–7.
- [45] Lo YM, Tein MS, Lau TK, Hainers CJ, Leung TN, Poon PM, et al. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am J Hum Genet* 1998;62:768–75.
- [46] Lun FMF, Chiu RWK, Chan AKC, Lau TK, Leung TY, Lo DYM. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin Chem* 2008;54:1664–72.
- [47] Higgs DR. In: Steinberg MH, Forget BG, Weatherall DJ, editors. *Disorders of hemoglobin*. 2nd ed. Cambridge: Cambridge University Press; 2008.
- [48] Higgs DR, Lamb J, Aldridge BE, et al. Inadequacy of Hb Bart's as an indicator of alpha thalassemia. *Br J Haematol* 1982;51:177–8.
- [49] Laosombat V, Viprakasit V, Chotsampancharoen T, Wongchanchailert M, Khodchawan S, Chinchang W, et al. Clinical features and molecular analysis in Thai patients with HbH disease. *Ann Hematol* 2009;88:1185–92.
- [50] Higgs DR, Weatherall DJ. The alpha thalassaemias. *Cell Mol Life Sci* 2007;66:1154–62.
- [51] Arnold SC, Quah TK, Low PS, Chong SS. A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for  $\alpha$ -thalassaemia. *Blood* 2001;98:250–1.



- [52] Garewal G, Fearon CW, Warren TC, Marwaha N, Marwaha RK, Mahadik C, et al. The molecular basis of  $\beta$  thalassemia in Punjabi and Maharashtrian Indians includes a multilocus aetiology involving replicated  $\alpha$ -globin loci. *Br J Haematol* 1994;86:372–6.
- [53] Nadkarni A, Gorashaker AC, Lu CY, Krishnamurthy R, Ghosh K, Colah R, et al. Molecular pathogenesis and clinical variability of  $\beta$ -thalassemia syndromes among Indians. *Am J Hematol* 2001;68:75–80.
- [54] Trehan U, Garewal G, Kaul D, Das R. Alpha thalassemia and alpha gene triplications in Punjabis, with and without beta thalassemia trait. *Haematology* 2000;6:153–60.
- [55] Shaji RV, Eunice SE, Baidya S, Srivastava A, Chandy M. Determination of the breakpoint and molecular diagnosis of a common alpha-thalassaemia-1 deletion in the Indian population. *Br J Haematol* 2003;123:942–7.
- [56] Jamwal M, Aggarwal A, Kumar V, Sharma P, Sachdeva MU, Bansal D, et al. Disease-modifying influences of coexistent G6PD-deficiency, Gilbert syndrome and deletional alpha thalassemia in hereditary spherocytosis: a report of three cases. *Clin Chim Acta* 2016;458:51–4.
- [57] Gupta RB, Tiwary RS, Pande PL, Kutlar F, Oner C, Oner R, et al. Hemoglobinopathies among the Gond tribal groups of central India; interaction of alpha- and beta-thalassemia with beta chain variants. *Hemoglobin* 1991;15(5):441–58.
- [58] Wajcman H, Traeger-Synodinos J, Papassotiropoulos I, Giordano PC, Cornelis L, Hartevelde CL, et al. Unstable and thalassaemic  $\alpha$  chain hemoglobin variants: a cause of Hb H disease and thalassemia intermedia. *Hemoglobin* 2008;32(4):327–49.
- [59] Wang W, Ma ESK, Chan AYY, Prior J, Erber WN, Chan LC, et al. Single-tube multiplex-PCR screen for anti-3.7 and anti-4.2  $\alpha$ -globin gene triplications. *Clin Chem* 2003;49:1679–82.
- [60] Ho PJ, Hall GW, Luo LY, Weatherall DJ, Thein SL. Beta-thalassemia intermedia: is it possible consistently to predict phenotype from genotype? *Br J Haematol* 1998;100:70–8.
- [61] Camaschella C, Mazza U, Roetto A, Gottardi E, Parziale A, Travi M, et al. Genetic interactions in thalassemia intermedia: analysis of beta-mutations, alpha-genotype, gamma-promoters, and beta-LCR hypersensitive sites 2 and 4 in Italian patients. *Am J Hematol* 1995;48:82–7.
- [62] Weatherall DJ, Clegg JB. The thalassaemia syndromes. 4th ed. Oxford: Blackwell Science; 2001.
- [63] Rani N, Jamwal M, Kaur J, Sharma P, Malhotra P, Maitra A, et al. Homozygous KLF1 mutation c.901C > T (p. Arg301Cys) resulting in mild thalassemia intermedia in an Indian: a next-generation sequencing diagnosis. *Blood Cells Mol Dis* 2018;72:19–21.
- [64] Sreedharanunni S, Chhabra S, Hira JK, Bansal D, Sharma P, Das R.  $\beta$ -Thalassemia intermedia caused by compound heterozygosity for Hb lepre-hollandia and  $\beta$ -thalassemia is rare in the Indian population. *Hemoglobin* 2015;39(5):362–5.
- [65] Badyal RK, Chhabra S, Sharma P, Das R. An intriguing high performance liquid chromatogram of a double heterozygosity for Hb Q-India/Hb D-Punjab. *Hemoglobin* 2014;38:440–3.
- [66] Sharma S, Sharma P, Das R, Chhabra S, Hira JK. HbQ-India (HBA1:c.193G > C): hematological profiles and unique CE-HPLC findings of potential diagnostic utility in 65 cases. *Ann Hematol* 2017;96:1227–9.
- [67] Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Dewi M, et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *Lancet* 2013;381:142–51.
- [68] Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. *Lancet* 2010;376:2018–31.
- [69] Kumar M, Sharma P, Kumar V, Bhargava M. Continuing diagnostic relevance of the sickling test in the era of CE-HPLC. *Indian J Hematol Blood Transfus* 2013;29:58–60.
- [70] Bender MA. Sickle cell disease. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews®*. Seattle (WA): University of Washington, Seattle; 2003. p. 1993–2018. Sep 15 [updated 2017 Aug 17].
- [71] Fucharoen S, Weatherall DJ. The hemoglobin E thalassemias. *Cold Spring Harb Perspect Med* 2012;2(8). Available from: <https://doi.org/10.1101/cshperspect.a011734>.
- [72] HbVar: a database of human hemoglobin variants and thalassemias [cited 2014 Aug 04]. Available from: <http://globin.bx.psu.edu/cgi-bin/hbvar/counter>; 2014.
- [73] Kumar GV, Sharma P, Chhabra S, Hira JK, Trehan A, Das R. Hb M-Iwate in an Indian family. *Clin Chim Acta* 2015;446:192–4.
- [74] Li YQ, Ye LH, Mo Y. Detection of the unstable Hb Köln (HBB: c.295G > A) by a capillary electrophoresis method. *Hemoglobin* 2016;40:17–19.
- [75] Dash S, Das R. Late emergence of polycythemia in a case of Hb Chandigarh [ $\beta$ 94(FG1)Asp--> Gly]. *Hemoglobin* 2004;28:273–4.
- [76] Dailey HA, Meissner PN. Erythroid heme biosynthesis and its disorders. *Cold Spring Harb Perspect Med* 2013;3:a011676.

## Further reading

- Das R, Sharma P. Molecular genetics of thalassemia syndromes. In: Kumar D, editor. *Colloquium series – digital library of genomic and molecular medicine*, 5(1). Williston, ND: Morgan and Claypool Publishers; 2016. p. 1–57.
- Garewal G, Das R, Kaur J, Marwaha RK, Gupta I. Establishment of prenatal diagnosis for  $\beta$ -thalassemia: a step towards its control in a developing country. *Ann Hum Biol* 2005;32:138–44.
- Globin Gene Server. Laboratories of computer science, engineering, biochemistry and molecular biology. Pennsylvania State University <http://globin.cse.psu.edu>, 2016 (accessed 07.07.16).
- Weatherall DJ. *Pathophysiology of thalassaemia*. Baillieres Clin Haematol 1998;11:127–46.
- Wintrobe MM, Greer JP. *Wintrobe's clinical hematology*. 12th ed. Philadelphia, PA/London: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2009.





# Coagulation and bleeding disorders

*Indu Singh<sup>1</sup> and Raghu Inder Singh<sup>2</sup>*

<sup>1</sup>School of Medical Science, Griffith University, Southport, QLD, Australia <sup>2</sup>RP WELLTAR Hospital, Bastara, Karnal, India

## 19.1 Introduction

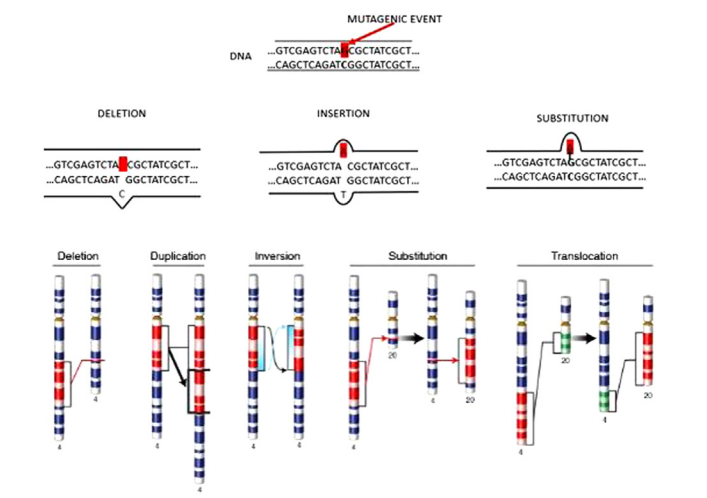
Thrombosis is a complex disorder involving interaction between many different types of genes with environment factors. Modification in genes coding for coagulation and regulatory proteins of hemostasis are important risk factors for thrombosis. Historically, the measurement of a clot endpoint has been the basis of coagulation testing mainly focused on bleeding disorders. Thrombosis diagnosis has slowly increased with most modern automated coagulation instruments utilizing optic, immunologic, and chromogenic methods. However, molecular diagnostics has added another dimension for the evaluation of clotting and bleeding disorders. DNA-based tests are commonly available for detection of the factor V Leiden (FVL) mutation, the prothrombin 20120A mutation, and the methyltetrahydrofolate reductase (MTHFR) mutation.

Inherited disorders of primary hemostasis include von Willebrand disease (VWD), congenital thrombocytopenia, Bernard–Soulier syndrome, Glanzmann thrombasthenia, and storage pool deficiencies. Secondary hemostasis congenital disorders include hemophilia A (factor VIII deficiency), B (factor IX deficiency), C (factor XI deficiency), other factor deficiencies, and dysfibrinogenemia. Thrombus dislodged from the blood vessel and moving through vasculature as an embolus plays a crucial role in the pathogenesis of acute myocardial infarctions, stroke, and venous thrombosis often leading to fatality.

For more than two decades now, FVL [1] and G20210A mutations in the nontranslational region 3' of the prothrombin gene [2] have been established as thrombotic risk factors in about 50% of the global population. However, in Asia and subcontinent of India, prevalence of these genetic aberrations seems to be much lower. It seems that single-nucleotide polymorphisms (SNPs) within or close to some hemostasis-associated candidate genes may be responsible for thrombosis and bleeding disorders. Some of the mutations, such as A384S in SERPINAC1 gene [3], coding for antithrombin (AT), and R67X mutation in SERPINC gene [4] coding for protein Z inhibitor [5], have also been associated with increased risk of thrombosis. It has also been shown that blood group A1 allele is also associated with increased risk of thrombotic events [6–8].

Homologous areas of the genomic DNA contain variations within their nucleotide sequence. When this sequence variation has a greater than 1% frequency in the population, it is referred to as a polymorphism [9]. These polymorphisms can occur in genes and gene-related sequences or noncoding extragenic DNA. Alleles represent polymorphisms within a single gene [10]. Fig. 19.1 shows various forms of changes in DNA sequences resulting in mutation.

When the sequence variation has a population frequency of less than 1%, this might be referred to as a mutation; however, some mutations might be present with >1% frequency (e.g., FVL). Many different polymorphisms/mutations form the molecular basis of risk factors associated with coagulation disorders. With several genes of coagulation proteins now completely sequenced, several SNPs have been described, including point mutations in exons, introns, or regulatory regions [11]. These mutations can result in loss of function (e.g., protein C deficiency) or gain of function (e.g., prothrombin mutation 20210A) [11,12]. A single mutation in the F5 gene is



**FIGURE 19.1** The figure shows various forms of changes in DNA sequences resulting in mutation. Adapted from National Institutes of Health. National Human Genome Research Institute. Talking glossary of genetic terms. Retrieved August 30, 2018, from <https://www.genome.gov/glossary/>.

associated with resistance to activated protein C thus leading to thrombotic tendency. However, there is a remarkable difference in the prevalence of these mutations in individuals of distinct ethnic backgrounds.

Different mutations and polymorphisms might affect the same gene; there might be silent mutations or frame-shift mutations, the latter typically associated with a severe phenotypic abnormality [11]. Levels of coagulation factors might vary depending on the type of polymorphism or mutations and their interaction with each other [12,13]. This degree of genetic complexity remains a challenge for the molecular diagnostics laboratory.

It has been established for a long time that hemophilia (bleeding disorders) has a genetic predisposition. Classic hemophilia (i.e., A and B) arises from hundreds of distinct mutations in the F8 genes, whereas 40% of severe hemophilia A results from a unique mutation characterized as F8 gene inversion. Genetic characterization of hemophilia A and B, the X-linked disorders, are now routinely incorporated into the standard of care, and genetic information is used for risk stratification of treatment complications. With electronic databases detailing >2100 unique mutations for hemophilia A and >1100 mutations for hemophilia B, these diseases are among the most extensively characterized inherited diseases in humans [14]. Mutations in some genes expressing coagulation proteins can lead to rare bleeding disorders including prothrombin, FV, FVII, FX, FXI, FXIII, combined FV and FVIII, or vitamin K-dependent factor deficiency.

Molecular diagnostics plays an important role in the clinical management of inherited bleeding disorders. Identification of mutation in bleeding disorders, in particular hemophilia, is crucial for carrier and prenatal testing as well as prediction of risk for inhibitor formation or anaphylactic reaction after replacement therapy [15]. Molecular diagnostics is also important for differentiating the bleeding disorder with similar clinical presentations but different underlying genetic causes. Informed decision of treatment and clinical management requires correct diagnoses of any genetic condition resulting in bleeding.

In most cases, initial diagnosis is made by traditional coagulation-based assays. Molecular testing provides additional details to aid in the future clinical management of these disorders.

Hemophilia B Leiden phenotype has now been extensively characterized at the genetic level and found to be caused by a group of single-nucleotide substitutions clustered around the transcription start site of the F9 gene [16].

Mild or moderate hemophilia A has been sometimes underdiagnosed by traditional coagulation testing. It has been revealed that the cause of this phenotype is missense substitutions localized to the interfaces between the factor VIII "A" domains that result in increased instability of the molecule [17].

Conventional coagulation-based testing could be sometimes nonspecific and insensitive to mild factor deficiencies [18]. Tests of intermediate complexity, such as factor assays, platelet aggregometry, and flow cytometry, are dependent on preanalytical variables [19]. On the other hand, heterogeneity of molecular lesions makes it technically complex to use them as first-line diagnostic tests. Current applications of molecular testing in these disorders are therefore largely restricted to carrier detection and antenatal diagnosis [20]. Later generation—sequencing techniques such as multiplexing have overcome many of the shortcomings of the conventional molecular assays with

improved sensitivity and lower costs. Molecular testing provides additional details to aid in the future clinical management of these disorders. Molecular diagnostic testing routinely performed include in vitro testing for nucleic acids by polymerase chain reaction (PCR) and reverse transcriptase PCR (RT-PCR). DNA and RNA extractions followed by DNA amplification is then detected using either gel electrophoresis, restriction endonucleases, nucleic acid hybridization or southern blotting cleavage-based signal amplification and DNA sequencing.

This chapter will introduce molecular systems underpinning various thrombotic and bleeding disorders in clinical setting with one case study description each for thrombosis and bleeding disorder.

## 19.2 Genetic basis of thrombosis

### 19.2.1 Case 1

A young lady of 25 years came to the doctor's clinic with deep vein thrombosis (DVT) of left lower limb. She was treated for that and then she returned after 6 months with DVT symptoms in her right lower limb. Her coagulation profile was completed confirming antiphospholipid syndrome (APLS). She returned in her first trimester of pregnancy after 2 years with blue discoloration that seemed like thrombosis of left great toe. She was referred to a tertiary specialist institution, where she was started on low molecular weight heparin (LMWH). However, she aborted at about 20 weeks. Thereafter she had three pregnancies which all ended in abortion. About a year ago, she again came back with severe pain in her abdomen, which was finally diagnosed as superior mesenteric artery thrombosis. Is it possible that it is some gene mutation or inherited genetic abnormality? Could genetic testing have helped in this case?

### 19.2.2 Molecular genetics of APLS

Human leukocyte antigen (HLA) class II antigens (DR, DP, DQ) loci are found on chromosome 6, and these molecules are highly polymorphic. HLA class II polymorphisms have been associated with autoimmune responses. Another gene implicated is  $\beta 2$ -GPI gene located on chromosome 17 with at least 4 common SNPs in protein coding region. It has also been reported that APLS population or those with APL antibodies had higher incidence of FV mutation in patients presenting with venous thromboembolism (VTE). Another study reviewed by same authors reported an association between FVL, G20210A, in various combinations, with APLS-potentiated risk factor for venous thrombosis but had minimal risk of arterial thrombosis. It has also been demonstrated that antiphospholipid syndrome (APS) patients with MTHFR 677TT genotype had a lower mean age at first thrombotic event as was the situation with the patient in Case 1 earlier. On the contrary, other studies did not find significant association between VTE and APL. Since several genetic and ethnicity factors are involved in pathophysiology of APS, thrombosis and recurrent pregnancy loss, internationally collaborated multicenter case-control studies are required for better understanding of genetic predisposition to develop APLS and patient management [21].

Sharma et al. reported APC resistance (APCR), followed by PS, PC, and AT deficiency, to be most common in Indian women presenting with recurrent pregnancy loss. They found that APCR associated with FVL increased risk of pregnancy loss in the population they investigated [22]. However, they believe their study is the only one conducted in India at the time of this publication, so there is scant data to make a conclusive statement.

### 19.2.3 Molecular basis of other causes of thrombosis

#### 19.2.3.1 Antithrombin deficiency

ATIII is a protein expressed by SERPINC1 gene spanning 13.4 kb of genomic DNA on chromosome 1q23–25. It inhibits activity of thrombin, factors IXa, Xa, and XIa. ATIII deficiency can be Type 1 where both activity and plasma antigen levels are attenuated or Type 2 with normal antigen levels but reduced activity [23].

More than 250 mutations and polymorphisms in protein coding region of SERPINC1 gene are known to cause DVT, but until 5 years ago, only one SNP (g.2085T > C) was identified in the SERPINC1 regulatory region. Recently, Bhakuni et al. identified two more SNPs (g.25G > A and g.-1A > T) and two previously known (g.67G > A and rs3138521) polymorphisms in Indian patients presenting with DVT for first time [23].

### 19.2.3.2 Protein C and S deficiency and activated protein C (APC) resistance

Protein C pathway and its cofactor protein S inactivates coagulation cofactors Va and VIIIa, which enhance the activity of serine protease coagulation factors VII, IX, X, XI, and XII, thus slowing down the clot-formation process. Protein C is a single-chain vitamin K-dependent protein synthesized by liver. Protein C gene is 11 kb of DNA on chromosome 2 with 9 exons and 8 introns. Protein C is activated by cleavage at Arg 169–Leu 170 [24].

Mutations in *PROS1* gene cause protein S deficiency. Like protein C-deficient individuals, protein S-deficient people cannot inactivate clotting proteins, resulting in the increased risk of developing abnormal blood clots. Protein S deficiency can be divided into Types 1, 2, and 3 based on how mutations in the *PROS1* gene affect protein S levels and activity [24].

APCR is a common cause of hereditary thrombophilia. It results from a substitution of G to A base at nucleotide 1691 in exon 10 of F5 gene. APCR prolongs inactivation of activated FV allowing the activation of other downstream coagulation factors and continuing or enhancing clot formation with a risk of thrombophilia. APCR alone is not a significant risk factor but combined with other risk factors increases risk of thrombosis [24].

#### 19.2.3.2.1 Prothrombin allele G20210A

Factor II (prothrombin) G20210A is the third most common cause of cardiovascular disease (CVD) and the most common inherited coagulation disorder in the United States. With autosomal-dominant inheritance, this single point mutation (G to A at position 20210) in prothrombin gene affects 80 times increase in thrombosis in homozygous state. G to A translation at nucleotide 20210 in prothrombin gene leads to an increase in factor II (prothrombin) levels increasing the risk of venous thrombosis.

The most common testing protocol to detect this mutation is PCR coupled with restriction-endonuclease digestion, gel electrophoresis, and RT-PCR [24]. Challenging clinical issues include the decisions regarding when to test for the mutation and how to manage individuals with the mutation, either in the setting of VTE or as an incidental finding. It has been used for patients with clinically suspected thrombophilia. There may be additional indications for direct PT G20210A mutation testing, such as in determining the duration of anticoagulation therapy of VTE patients and screening for women contemplating hormone therapy.

### 19.2.3.3 Factor V Leiden

A specific mutation in the factor V gene is called FVL. This mutation causes factor V to be inactivated more slowly by APC, generating more thrombin and consequently increasing the potential for clot formation. It is inherited in an autosomal dominant manner. Factor V is a plasma glycoprotein of 330 kDa molecular weight with 2224 amino acids. It is coded by a complex 25 exons—80 kb gene on chromosome 1. It is converted to an active two-chain form by thrombin or factor Xa. Thrombin cleaves it at three separate sites. The APC cleavage sites are 306, 506, and 679 on an FVa heavy chain, which is hinged by  $\text{Ca}^{2+}$  to the FVa light chain. Following the cleavage, 2 chains are linked via a divalent metal ion bridge.

FV acts as a cofactor for APC and protein S in the inactivation of the procoagulant FVIIIa. A cleavage of FV at R506 by APC is required for this cofactor function. FVL, the mutant FV, is seen with APCR. In this mutant, R506 is replaced by a Q, which renders the 506-position insensitive to proteolysis by APC. This is referred to as Q506-FV which does not show anticoagulant activity [25].

Inactivation of mutant FVa: Q506 site leads to FVL. FVL mutation accounts for about 90% of APCR and is prevalent in nearly 2%–13% of general population. It is also seen in about 20%–60% of VTE cases. Literature shows that a higher proportion of Caucasians are affected than other populations. This could also be relative, due to lack of studies and data from Asian and other populations. About 1 in 10 people carrying FVL heterozygotes develop VTE over their lifetime. Many patients with familial thrombophilia are found to have FVL. Other coexisting disorders and circumstantial risk factors also affect clinical expression of FVL. These include presence of prothrombin variant G20210A, obesity, age, immobility due to injury, surgery, oral contraception, hormone replacement therapy, pregnancy, and even air travel [26]. It has been suggested and is advisable to offer genetic testing and referral for consultation in first unprovoked VTE at any age, history of recurrent VTE, strong family history of thrombotic diseases, VTE during pregnancy or puerperium or when it is the first VTE incident and patient has a first-degree family member with a history of VTE under the age of 50. Case 1 in this chapter is a good example where offering genetic testing might be useful for patient and future generations. It has been suggested that 10% of FVL carriers develop VTE in their lifetime, and 25%–40% are at higher risk where there is a family history of thrombophilia.



One of the potential benefits of genetic testing is that asymptomatic FVL carriers could be educated about VTE high-risk circumstances, signs and symptoms, as well as possible potential need for prophylactic treatment in high-risk circumstances. However, genetic testing does not detect all possible thrombophilia conditions and has limited clinical utility. Further research is required to confidently improve surveillance and management guidelines as there is very limited and inconsistent data, if at all, in use currently.

#### **19.2.3.4 Hyperhomocysteinemia**

MTHFR gene is responsible for production of enzyme MTHFR which has an important role in the homocysteine metabolism. MTHFR gene has at least two known polymorphisms, 677C > T and 1298A > C. These are associated with reduced enzyme activity, decreased folate concentration in blood, mildly increasing plasma, and total homocysteine concentration [24].

#### **19.2.3.5 Inherited deficiency of fibrinolysis**

Decreased fibrinolytic capacity due to increased plasminogen activator inhibitor-1 (PAI-1) activity and decreased tissue-type plasminogen activator (tPA) activity has been associated with thrombosis. A congenital gamma-chain molecular defect or gamma dysfibrinogenemia such as fibrinogen Dusard (Arg554Cys) results in impaired binding of tPA to fibrin-reducing plasminogen activation, impaired fibrinolysis, and tendency for thrombosis [27]. PAI-1 gene has also been shown to have multiple polymorphisms. However, only patients with 4G/4G genotype of the PAI-1 genetic polymorphism show an increased plasma PAI-1 concentration which is significantly associated with an increased risk of thrombosis [28]. Plasminogen is activated to plasmin by specific cleavage at the Arg560–Val561 peptide bond.

### **19.3 Genetic basis of the bleeding disorders**

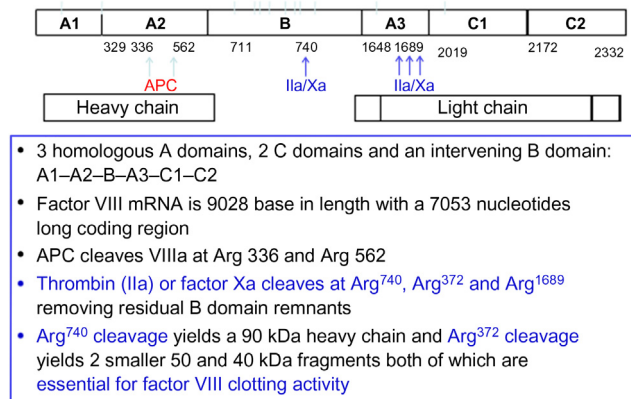
Bleeding disorders in men are well recognized due to a lot of research in X-linked recessive disorders, such as hemophilia, seen in young boys with prominent clinical significance. However, bleeding disorders in females are often missed due to milder presentation at later age. Some studies have focused on hemostatic disorders in menorrhagia [29,30]. Gupta et al. demonstrated that inherited bleeding disorders were present in most of the 200 women with various bleeding complaints in a tertiary hospital in India [31]. Current testing method, widely used by many accredited research-based facilities worldwide, includes Sanger or next-generation sequencing (NGS) for the testing of single-nucleotide variants particularly in hemophilia [32].

#### **19.3.1 Case 2**

A 14-year-old boy who had repeated episodes of hemarthrosis affecting his knees ultimately ended with flexion deformity of his right knee. The coagulation and factor testing reported factor VIII deficiency. It seems to be an obvious case of hemophilia, but do genetic and molecular studies on this patient have any role in patient care or future counseling? What other bleeding disorders could be suspected? There is no clear discrimination between inherited or acquired mutations leading to bleeding disorders. The complexity of structure with variable gene expression as well as degree of consanguinity in certain populations, concentrating effect of certain mutations within community, indicate that there may be a benefit in investigating patient's history, and molecular as well as cytogenetic evaluation to rule out coexisting recessively inherited risk factors.

#### **19.3.2 Hemophilia A (factor VIII deficiency)**

Hemophilia A is an X-linked recessive disorder with factor VIII deficiency mostly in male population. Rarely females may be presented with this condition due to the inheritance of homozygous or compound heterozygous alleles or chromosomal abnormalities. Hemophilia A is caused by pathogenic variant in F8 gene encoding for FVIII protein. The F8 gene, that is 187 kb in size, comprises 25 introns and 26 exons, located on the long arm of the X-chromosome at the most distal band (Xq28) [33]. Ankala et al. found that intron 22 inversion in F8 gene is the most common cause of hemophilia A in Indian population followed by some cases of intron 1 inversion on the same gene [34].



**FIGURE 19.2** Factor VIII molecule and its cleavage sites.

Most pathogenic hemophilia A variants reduce the synthesis or secretion of FVIII or impair FVIII cofactor activity. A small subset of variants, localized to the A3, C1, and C2 domains (Fig. 19.2), are associated with impaired binding to the FVIII carrier molecule von Willebrand factor (VWF) [35], resulting in accelerated proteolysis and/or clearance of FVIII from the plasma. More than 2000 unique F8 gene variants have been reported in the FVIII-variant database [36].

The most common mutation in severe hemophilia A is inversion, producing abnormal or reduced amount of this protein. About half of the hemophilia A cases have reported an intrachromosomal inversion at intron 22 occurring spontaneously during meiosis in male germ cells [37] and a similar inversion has been reported within intron 1 accounting for about 1%–2% cases only [38].

Most of the other heterogeneous cases are a result of nonsense/stop mutation preventing factor VIII production, missense mutations' impact on activity, half-life or efficient factor VIII production, or may be due to some new mutations.

Mutations in hemophilia A patients can be detected by chemical cleavage of mismatch method. Inversion mutations can be analyzed by restriction enzyme analysis and southern blotting. Molecular techniques including inverse-shifting PCR, long-range PCR, eight multiplex PCRs, and Sanger sequencing are also useful in diagnosis of these cases.

Widely used FVIII-replacement therapy has the risk of FVIII-neutralizing antibodies or inhibitor development in patient. The genetic variability within the F8 gene as well as environmental and genetic risk factors may be able to regulate immune response to FVIII. Intron 22 inversion is associated with lower risk for inhibitors. FVIII mutation status can help predict patient's response to immune tolerance induction (ITI) therapy for inhibitor eradication. Some large deletions have been shown to be associated with a decreased rate of ITI success [39].

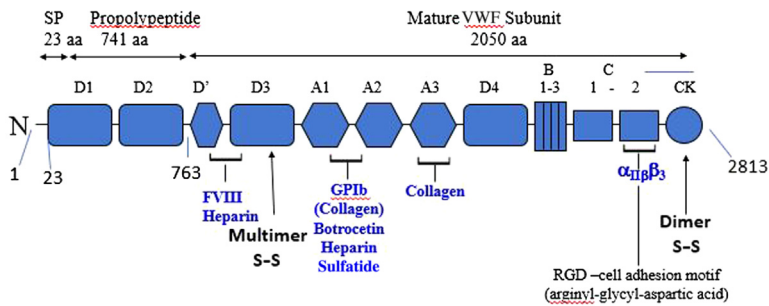
### 19.3.3 Molecular basis of other inherited coagulopathies

#### 19.3.3.1 von Willebrand disease

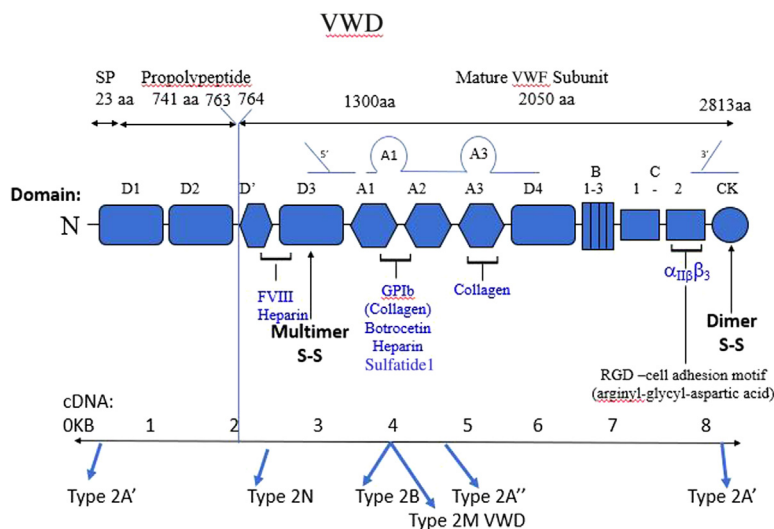
VWD is the most prevalent of the congenital bleeding disorders affecting both sexes through autosomal-dominant inheritance. It is a highly heterogeneous disorder due to the molecular mechanisms that produce various clinical and laboratory phenotypes. VWF gene is located on the short arm of chromosome 12 in the locus p13.3. It consists of 52 exons spanning 178 kilobase pairs [40].

VWF gene encodes for the VWF synthesized in endothelial cells and megakaryocytes. There are patients presented with excessive mucocutaneous bleeding caused by quantitative or qualitative abnormalities in VWF. VWF is a multimedia glycoprotein essential for the maintenance of hemostasis.

Multidomain structure is composed of multiples of four domain types A–D in the arrangement D1–D2–D'–D3–A1–A2–A3–D4–B1–B2–B3–C1–C2 (Fig. 19.3). Different functions of VWF are assigned to different domains: A1 domain is involved in binding of VWF to platelet GP1b binding to fibrillar collagen, sulfatides, and heparin. D3 domain binds to factor VIII. C2 domain binds to GP11b-111a receptors. After initial synthesis, it is modified by posttranslational processing of dimerization and multimerization by cleavage of D1–D2 domain of the polypeptide.



**FIGURE 19.3** VWF domain structure. VWF, von Willebrand factor.



**FIGURE 19.4** Most type 1 VWD are due to missense mutations (dominant negative—interference with intracellular transport of dimeric pro-VWF). Some forms with incomplete penetrance require coinheritance of blood type O for expression (causes increased VWD proteolysis). Most type 3 VWD due to null alleles. VWD, von Willebrand disease; VWF, von Willebrand factor.

VWD is a result of either quantitative deficiency (Type 1), qualitative abnormality (Type 2), or complete absence of VWF (Type 3). Mutations giving rise to Type 2 are very well characterized, whereas those giving rise to Type 1 are not well characterized. Type 2A with point mutation in A2 structural domain of VWF renders VWF more susceptible to proteolysis leaving a lot of small molecular weight multimers in plasma. Type 2B is a rare mutation in A1 structural domain of VWF reducing large molecular weight multimers due to increased binding with resting platelets. Type 2M is a qualitative variant of VWF reducing normal multimers due to high platelet receptor binding. Type 2N results from auto recessive-translation or deletion mutations of VWF gene, where VWF and factor VIII are nearly absent (Fig. 19.4).

People with VWF levels at upper or lower end of normal reference range are at risk of hemostatic disorders. Lower levels may lead to the common bleeding disorder Type 1 VWD, and high levels are associated with an increased risk for both venous and arterial thrombosis [41,42].

Most of the Type 1 VWD individuals with VWF levels under 40 have VWF missense gene mutation. Type 2 qualitative defect is often a result of missense mutation including Type 2A with mutations in either of the two domains involved in multimer formation or mutation in domain cleaved by ADAMTS-13. Type 2B shows gain of function mutation in platelet GP1b-binding domain, while Type 2M is a result of the loss of function mutation in GP1b-binding domain and Type 3 patients with severe deficiency are homozygous for null alleles.

Results from genome-wide association studies accounted for common variants at *ABO*, *VWF*, and other loci in about 12% of the variance in plasma VWF levels [43]. The genetic defects responsible for VWD include a novel gene on chromosome 2 accounting for about 1 in 5 of VWF variations not identified by standard genetic approaches. Further characterization of this and other genes controlling VWF levels may lead to improved VWD diagnosis and prediction of bleeding and clotting risk. In Indian population, mutations c.2908delC and c.5335C > T (p.R1779\*) on VWF gene have been found to cause VWD [44].

### 19.3.3.2 Hemophilia B (factor IX deficiency)

Hemophilia B is a quantitative or qualitative defect of the coagulation zymogen factor IX presented as an X-linked recessive disorder. The incidence is lower than hemophilia A. Factor IX gene is located on the long arm of X-chromosome (Xq27) and has eight exons. The outcome of the review of literature by Ankala et al. suggested that mutation c.316G > A (p.G106S) is the most common mutation of F9 gene in Indian population [34].

Heterogenous pathogenic variants are found throughout the factor IX gene including the promoter and 3'UTR region. Over 1000 unique variants, majority of them being single-nucleotide substitutions, have been associated with hemophilia B [36]. Most of these variants are on exon 8, the largest exon encoding the functional serine protease domain of the active factor IX coagulation protein.

Hemophilia B Leyden is found in about 2% cases presents with moderate to severe prepubertal factor IX deficiency, which eventually returns to normal. This is associated with variants in proximal promoter region of F9 gene and disrupts binding sites to one of the three transcription factors: hepatic nuclear factor 4a, CCAAT enhancer-binding protein (C/EBP), or ONECUT1/2 [45–47]. Sequencing the promoter proximal region of F9 gene can be predictive of long-term normalization of factor IX levels.

The incidence of inhibitors in these patients is significantly lower than hemophilia A cases. Chitlur et al. reported that about 60% of hemophilia B patients that develop inhibitors often experience an anaphylactic response to FIX concentrate at the time of inhibitor onset. They believe that this is more likely associated with large deletions or null mutations rather than single-nucleotide variants [48].

### 19.3.3.3 Afibrinogenemia and dysfibrinogenemia

Most cases of reduced or ineffective activity of fibrinogen have been reported from consanguineous parents. Inherited in recessive manner, these disorders are caused by more than 30 genetically heterogenous mutations. It results in moderate to severe bleeding, which could be due to low or zero synthesis of fibrinogen or problems with intracellular transport or secretion. Dysfibrinogenemia usually exhibits dominant inheritance caused by missense mutations affecting fibrin polymerization, fibrinopeptide cleavage, or fibrin stabilization by factor XIIIa. Depending on the mutation, the patient may have mild to severe bleeding tendency.

## 19.4 Structural defects of the vascular system

### 19.4.1 Hereditary hemorrhagic telangiectasia

This condition has an autosomal dominant inheritance caused by mutation in endoglin gene that controls vascular remodeling. Molecular diagnosis is possible for this condition showing several small arteriovenous malformations in skin, mouth, GI tract, and lungs [24].

### 19.4.2 Ehlers–Danlos syndrome

Ehlers–Danlos syndrome (EDS) is a defective collagen structure affecting connective tissues supporting the skin, bones, blood vessels, and many other organs and tissues. EDS is caused by the mutations in more than a dozen genes for various types of collagen with nine different variants, which leads to bruising due to thin, weak skin with poor healing and spontaneous joint dislocation due to hypermobile joints. Inheritance can be autosomal, dominant, recessive, or X-linked [24]. In 2017 EDS was classified describing 13 types. The specific gene affected determines the specific EDS in a patient. Till date, about 18 genes with mutations responsible for EDS have been identified. Some of these are fibrous protein genes, such as COL1A1, COL1A2, COL3A1, COL5A1, and TNXB, and enzymes including ADAMTS2, PLOD1, B4GALT7, DSE, and D4ST1/CHST14. Customized NGS is performed using a panel of various collagen genes. Analysis includes the coding exons of all genes in the panel plus 10 bases in the introns and untranslated regions (5' and 3'). Sanger sequencing is performed to confirm variants suspected or confirmed to be pathogenic [49]. However, there is no known cure for any type of EDS, so treatment is supportive rather than therapeutic.

## 19.5 Inherited defects of platelets

### 19.5.1 Inherited macrothrombocytopenia

Giant platelets seen in peripheral blood could be due to genetic mutations or polymorphism leading to inherited platelet defects. Recent evidence from the literature and testing of different populations from various areas of India [50,51] focused on the presence of inherited disorder where patients have high number of giant platelets. They found an increased recognition of inherited macrothrombocytopenia in the country. Their findings suggest that about half of the patient population showed mutations of MYH9, GP1BB, GP1BA, GPIX, ABCG5, ABCG8, ACTN, FLI, TUBB, and RUNX1 [50], which are frequently seen in heterozygous state. Many of these asymptomatic patients do have mild-to-moderate bleeding history.

### 19.5.2 Bernard–Soulier syndrome

Bernard–Soulier disease is caused due to reduced or absent levels of platelet glycoproteins GPIb/IX/V, which is one of the VWF receptors on platelets. This VWF receptor on platelets has four polypeptide chains: GP1b- $\alpha$ , GP1b- $\beta$ , GPV, and GPIX. Mutations in GP1BA or GP1BB genes reduce expression of these genes. GP1BA insertion mutations (p.Met338fsX13) resulting in frameshift change and (p.Val485fsX13) a new termination codon were reported in people from India. Two missenses (p.Tyr95Asp and p.Cys32X), also resulting in frameshift, as well as one missense (p.Cys24Arg) in GPIX mutation, were also found in some Indian patients [52]. However, this study was conducted on a very small number of eight patients from seven unrelated families with six different detected mutations. This may not have the power to make a generalization about any set population, but it does show high variability, adding further to the complexity of using molecular diagnostics as a tool in many cases.

### 19.5.3 Glanzmann's thrombasthenia

Autosomal recessive inheritance with decreased platelet GPIIb–IIIa expression resulting in defective platelet aggregation and moderate to severe bleeding. It is common in populations with high degree of consanguinity. More than 70 mutations in  $\alpha_{IIb}$  and  $\beta_1$  genes on chromosome 17 have been found to be responsible for this disorder [53].

### 19.5.4 Storage pool disease

Dense and alpha granule deficiency is caused by variety of genetic abnormalities. Mutations in a gene for a nucleotide transporter MRP4 (ABCC4) have been suggested to affect nucleotide accumulation in dense granules [54].

### 19.5.5 May–Hegglin anomaly

It is a giant platelet syndrome associated with mutations in the nonmuscle myosin heavy chain gene MYH9 with autosomal dominant inheritance [24].

### 19.5.6 Wiscott–Aldrich syndrome

Mutation in WASP-signaling protein decreases secretion and aggregation of platelets with multiple agonists. It has X-linked inheritance. Current diagnosis is based on family history, clinical thrombocytopenia, small platelets with few granules, and genetic testing [24].

## 19.6 Conclusion

Thrombophilia is a very complex multifactorial disorder. However, so far there are less than 10 well-documented genetically associated risk factors of thrombosis. These defects account for a very low percentage of hereditary thrombophilia and thrombosis leading to a very common and much higher prevalence of



thrombosis-related cardiovascular events worldwide. Considering these known genetic factors have low frequency particularly in Asia with highest population concentration, there is a need to identify new genetic factors contributing to variation as a collaborative effort amongst different ethnic groups instead of fragmented projects investigating random patient populations. This knowledge will help us not only to personalize the therapeutic algorithms but to propose preventive strategies. There are ever evolving methods, automation, and statistical tools for genetic testing and interpretation to guide effective diagnosis, prophylactic, and therapeutic strategies. Rapid advances in molecular genetics have informed current approaches to the diagnosis of many hemostatic and thrombotic disorders and have revolutionized the treatment of patients with hemophilia with the introduction of recombinant clotting proteins. However, curative gene therapy for inherited disorders of coagulation remains elusive, although continued improvements in the tools for effective and safe gene therapy suggest a bright future for genetic therapies of hemophilia.

Dramatic advances in our understanding of the mechanisms of thrombosis and hemostasis have transformed the management of thrombotic and bleeding disorders through the introduction of new, effective, safer, and more convenient therapies. There are still lots of gaps with unknowns, and future studies need to identify directions of research and variation in different populations of world.

### Conflict of interest

There are no conflicts of interest to declare.

### Author's roles

The chapter was written by IS in consultation with RIS and was reviewed by RIS.

### References

- [1] Bertina RM, Koeleman BPC, Koster T, Rosendaal FR, Dirven RJ, De Ronde H, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 1994;369:64–7.
- [2] Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* 1996;88:3698–703.
- [3] Corral J, Hernández-Espinosa D, Soria JM, González-Conejero R, Ordoñez A, González-Porras JR, et al. Antithrombin Cambridge II (Ala384Ser): an underestimated but strong and frequent genetic risk factor for venous thrombosis. *Blood* 2007;115:4258–63.
- [4] Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT Study. *Am J Hum Genet* 2000;66:1452–9.
- [5] Corral J, González-Conejero R, Soria JM, González-Porras JR, Pérez-Ceballos E, Lecumberri R, et al. A nonsense polymorphism in the protein Z-dependent protease inhibitor increases the risk for venous thrombosis. *Blood* 2006;108:177–83.
- [6] Souto JC, Almasy L, Muniz-Diaz E, Soria JM, Borrell M, Bayen L, et al. Functional effects of the ABO locus polymorphism on plasma levels of von Willebrand factor, factor VIII, and activated partial thromboplastin time. *Arterioscler Thromb Vasc Biol* 2000;20:2024–8.
- [7] Morelli VM, De Visser MC, Vos HL, Bertina RM, Rosendaal FR. ABO blood group genotypes and the risk of venous thrombosis: effect of factor V Leiden. *J Thromb Haemost* 2005;3:183–5.
- [8] Wu O, Bayoumi N, Vickers MA, Clark P. ABO (H) blood groups and vascular disease: a systematic review and meta-analysis. *J Thromb Haemost* 2008;6:62–9.
- [9] Tefferi A, Wieben E, Dewald GW, Whiteman DAH, Bernard ME, Spelsberg TC. Primer on medical genomics: Part II: Background principles and methods on molecular genetics. *Mayo Clin Proc* 2002;77:785–808.
- [10] Rennert H, Leonard DGB. Basic principles of molecular biology. *Diagnostic molecular pathology*. Philadelphia, PA: W. B. Saunders; 2003. p. 25–51.
- [11] Reitsma PH. Genetic principles underlying disorders of procoagulant and anticoagulant proteins. In: Colman RW, Hirsh J, Marder VJ, Clowes AW, George JN, editors. *Hemostasis and thrombosis*. Philadelphia, PA: Lippincott Williams and Wilkins; 2001. p. 59–87.
- [12] Franco RF. Gene polymorphism of the haemostatic system and the risk of arterial thrombotic disease. *Br J Haematol* 2001;115:491–506.
- [13] Humphries SE, Panahloo A, Montgomery HE, Green F, Yudkin J. Gene environment interaction in the determination of levels of haemostatic levels involved in thrombosis and fibrinolysis. *Thromb Hemost* 1997;78:457–61.
- [14] Peyvandi F, Kunicki T, Lillicrap D. Genetic sequence analysis of inherited bleeding diseases. *Blood* 2013;122(20):3423–31.
- [15] Oldenburg J, Pavlova A. Genetic risk factors for inhibitors to factors VIII and IX. *Haemophilia* 2006;12:15–22.
- [16] Picketts DJ, Mueller CR, Lillicrap D. Transcriptional control of the factor IX gene: analysis of five cis-acting elements and the deleterious effects of naturally occurring haemophilia B Leyden mutations. *Blood* 1994;84(9):2992–3000.
- [17] Cid AR, Calabuig M, Cortina V, Casana P, Haya S, Moret A, et al. One-stage and chromogenic FVIII:C assay discrepancy in mild haemophilia A and the relationship with the mutation and bleeding phenotype. *Haemophilia* 2008;14(5):1049–54.

- [18] Toulon P, Eloit Y, Smahi M, Sigaud C, Jambou D, Fischer F, et al. In vitro sensitivity of different activated partial thromboplastin time reagents to mild clotting factor deficiencies. *Int J Lab Hematol* 2016;38:389–96.
- [19] Jury C, Nagai Y, Tatsumi N. Collection and handling of blood. In: Bain BJ, Bates I, Lewis JM, editors. *Dacie and lewis practical haematology*, 11. Maryland Heights, MO: Elsevier; 2012. p. 1–10.
- [20] Lillicrap D. Molecular testing for disorders of hemostasis. *Int J Lab Hematol* 2013;35:290–6.
- [21] Hancer VS. Genetics of antiphospholipid syndrome. *Huma Genet Embryol* 2011;1:e103. Available from: <https://doi.org/10.4172/2161-0436.1000e103>.
- [22] Sharma A, Bhakuni T, Ranjan R, Kumar R, Kishor K, Kamal VK, et al. Polymorphism in factor V and antithrombin III gene in recurrent pregnancy loss: a case-control study in Indian population. *J Thromb Thrombolysis* 2015;39:481–8.
- [23] Bhakuni T, Sharma A, Ali MF, Mahapatra M, Saxena R, Jairaipuri MA. Identification of 2 novel polymorphisms and rs3138521 in 5′ untranslated region of SERPINC1 gene in north Indian pupation with deep vein thrombosis. *Clin Appl Thromb Hemost* 2017;23(7):865–70.
- [24] Keohane E, Smith L, Walenga J. *Rodak's hematology: clinical principles and applications*. 5th ed. St. Louis, MO: Elsevier; 2018. Chapters 30–31 & 38–41.
- [25] Cramer TJ, Griffin JH, Gale AJ. Factor V is an anticoagulant cofactor for activated protein C during inactivation of factor Va. *Pathophysiol Haemost Thromb* 2010;37(1):17–23.
- [26] Kujovich JL. Factor V Leiden thrombophilia. *Genet Med* 2011;13(1):1–13.
- [27] Côté HC, Lord ST, Pratt KP. Gamma-Chain dysfibrinogenemias: molecular structure-function relationships of naturally occurring mutations in the gamma chain of human fibrinogen. *Blood* 1998;92(7):2195–212.
- [28] Lorente L, Martín MM, Borreguero-León JM, Barrios Y, Solé-Violán J, Ferreres J, et al. The 4G/4G genotype of PAI-1 polymorphism is associated with higher plasma PAI-1 concentrations and mortality in patients with severe sepsis. *PLoS One* 2015;10(6):e0129565. Available from: <https://doi.org/10.1371/journal.pone.0129565>.
- [29] Dilley A, Drews C, Miller C, Lally C, Austin H, Ramaswamy D, et al. von Willebrand's disease and other inherited bleeding disorders in women with diagnosed menorrhagia. *Obstet Gynecol* 2001;97:630–6.
- [30] Al-Zirqi I, Vangen S, Forsen L, Stray-Pedersen B. Prevalence and risk factors of severe obstetric haemorrhage. *BJOG* 2008;115:1265–72.
- [31] Gupta A, Mishra P, Pati HP, Tyagi S, Mahapatra M, Seth T, et al. Spectrum of hemostatic disorders in Indian females presenting with bleeding manifestations. *Int J Lab Hematol* 2018;40:437–41.
- [32] Gomez K, Chitlur M. Survey of laboratory tests used in the diagnosis and evaluation of haemophilia A. *Thromb Haemost* 2013;109:738–43.
- [33] Gitschier J, Wood WI, Goralka TM, Wion KL, Chen EY, Eaton DH, et al. Characterization of the human factor VIII gene. *Nature* 1984;312:326–30.
- [34] Ankala A, Tamhankar PM, Valencia CA, Rayan KK, Kumar MM, Hegde MR. Clinical applications and implications of common and founder mutations in Indian subpopulations. *Hum Mutat* 2015;36(1):1–10.
- [35] Jacquemin M, De Maeyer M, D'Oiron R, Lavend'Homme R, Peerlinck K, Saint-Remy JM. Molecular mechanisms of mild and moderate haemophilia A. *J Thromb Haemost* 2003;1:456–63.
- [36] Rallapalli P, Kemball-Cook G, Tuddenham E, Gomez K, Perkins S. Factor VIII variant database, <<http://www.factorviii-db.org>>; 2016.
- [37] Rossiter J, Young M, Kimberland M, Hutter P, Ketterling R, Gitschier J, et al. Factor VIII gene inversions causing severe haemophilia A, originate almost exclusively in male germ cells. *Hum Mol Genet* 1994;3:1035–9.
- [38] Bagnall RD, Waseem N, Green PM, Giannelli F. Recurrent inversion breaking intron1 of the factor VIII gene is a frequent cause of severe haemophilia A. *Blood* 2002;99:168–74.
- [39] Salvati R, Belvini D, Radossi P, Sartori R, Pierobon F, Zanotto D, et al. F8 gene mutation profile and ITT response in a cohort of Italian haemophilia A patients with inhibitors. *Haemophilia* 2007;13:361–72.
- [40] Lenting PJ, Casari C, Christophe OD, Denis CV. von Willebrand factor: the ld, the new and the unknown. *J Throm Haemost* 2012;10:2428–37.
- [41] Sadler JE. von Willebrand factor: two sides of a coin. *J Thromb Haemost* 2005;3:1702–9.
- [42] Sonneveld MA, de Maat MP, Leebeek FW. Von Willebrand factor and ADAMTS13 in arterial thrombosis: a systematic review and meta-analysis. *Blood Rev* 2014;8:167–78.
- [43] Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: the CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) consortium. *Circulation* 2010;121:1382–92.
- [44] Kasatkar P, Shetty S, Ghosh K. Genetic heterogeneity in a large cohort of Indian Type 3 von Willebrand disease patients. *PLoS One* 2014;9(3):e92575. Available from: <https://doi.org/10.1371/journal.pone.0092575>.
- [45] Crossley M, Brownlee GG. Disruption of a C/EBP binding site in the FIX promoter is associated with haemophilia B. *Nature* 1990;345:444–6.
- [46] Reijnen MJ, Sladek FM, Bertina RM, Reitsma PH. Disruption of a binding site for hepatocyte nuclear factor 4 results in hemophilia B Leyden. *Proc Natl Acad Sci USA* 1992;89:6300–3.
- [47] Funnell APW, Wilson MD, Ballester B, Mak KS, Burdach J, Magan N, et al. A CpG mutational hotspot in an ONECUT binding site accounts for the prevalent variant of haemophilia B Leyden. *Am J Hum Genet* 2013;92:460–7.
- [48] Chitlur M, Warrier I, Rajpurkar M, Lusher JM. Inhibitors in factor IX deficiency a report of the ISTH-SSC international FIX inhibitor registry (1997–2006). *Haemophilia* 2009;15:1027–31.
- [49] Genetics Home Reference. Ehlers-Danlos syndrome. Retrieved from <<https://ghr.nlm.nih.gov/condition/ehlers-danlos-syndrome>>; 2015.
- [50] Ghosh K, Bhattacharya M, Chowdhury R, Mishra K, Ghosh M. Inherited macrothrombocytopenia: correlating morphology, epidemiology, molecular pathology and clinical features. *Indian J Haematol Blood Transfus* 2018;34(3):387–97.

- [51] Patel P, Shah A, Mishra K, Ghosh K. Prevalence of macrothrombocytopenia in healthy college students in western India. *Indian J Haematol Blood Transfus* 2018;. Available from: <https://doi.org/10.1007/s12288-018-0970-9>.
- [52] Ali S, Ghosh K, Shetty S. Molecular pathology of Bernard-Soulier syndrome in Indian patients. *Platelets* 2013;24:571–3.
- [53] Nurden AT. Qualitative disorders of platelets and megakaryocytes. *J Throm Haemost* 2005;3:1773–82.
- [54] Jedlitschky G, Tirwschmann K, Lubenow LE, Nieuwenhuis HK, Akkerman JW, Greinacher A, et al. The nucleotide transporter MRP4 (ABCC4) is highly expressed in human platelets and present in dense granules, indicating a role in mediator storage. *Blood* 2004;104:3606–10.

## Molecular and genomic basis of bronchial asthma

Anubhuti Gupta<sup>1,\*</sup>, Samarpana Chakraborty<sup>1,\*</sup> and Anurag Agrawal<sup>1,2,3</sup>

<sup>1</sup>CSIR Institute of Genomics and Integrative Biology, Delhi, India <sup>2</sup>Academy of Scientific and Innovative Research, Delhi, India <sup>3</sup>Baylor College of Medicine, Houston, Texas, United States

### 20.1 Introduction

Bronchial asthma is an obstructive airway disease characterized by variable expiratory airflow limitation, bronchial hyperresponsiveness, and airway inflammation. Asthma exacerbation is usually an episodic response to allergens such as pollen, dust, animal hair, or external stimuli such as cold and exercise. Clinical signs and symptoms include cough, shortness of breath, dyspnea, wheezing, and chest tightness. Diagnosis of asthma is clinical as per the Global Initiative for Asthma guidelines. Spirometry is also used as an adjunct in diagnosis and is characterized by a decrease in forced expiratory volume in 1st second reversible by at least 12% and 200 mL after inhaling a bronchodilator [1].

Bronchial asthma is a complex genetic disease with an interplay of environmental exposures, genetic predisposition, epigenetic regulation, and host susceptibility leading to the clinical disease phenotype. The prevalence of asthma has been increasing worldwide with a rate of 1.8% yearly, and 300 million people are currently affected [2]. It is thus imperative to study the complex environment–gene interactions in asthma to find ways for understanding the biology behind asthma pathogenesis and better management of the disease.

### 20.2 Genomics of bronchial asthma

It has been known for decades that complex diseases, such as asthma, hay fever, and eczema, run into families. Testing and estimating familial aggregation of a disease consist of comparing rates of disease in relatives of individuals with the disease (known as case probands) with rates of disease in relatives of individuals without the disease (known as control probands). Relative risk (RR) is the prevalence of the disease in first-order relatives of affected individual compared to its prevalence in the general population. Higher the value of RR, greater is the contribution of genetics in the disease prevalence [3].

For asthma the general prevalence in a given population varies between 2% and 5%, except in developed nations where the prevalence is greater. But prevalence in affected families, on the other hand, ranges between 20% and 25%, which gives RR a value of about 5 or more. For other complex diseases, such as diabetes, the value of RR ranges between 3 (Type II diabetes) and 15 (Type I diabetes) [4]. This highlights the significant extent of the contribution of genetics in the case of asthma, in comparison to other complex diseases.

Further, heritability ( $h^2$ ) provides a measure of the relative importance of transmissible genetic effects in the overall phenotypic variation. Heritability for atopy and asthma has been estimated to be 0.4–0.6, that is, the relative contribution of genetic factors to atopy and asthma is estimated to be 40%–60% [5].

\* These authors contributed equally to this work.

Cooke and Vander [6], for the first time, undertook a comprehensive study to see if there was any consistent dominant or recessive mode of inheritance in asthma but failed to find so. This was expected as our current knowledge tells us that asthma being a complex disease follows no particular mode of inheritance. In 1959 Van Arsdel and Motulsky performed an epidemiological study on 6000 students of Washington's university, which further reinforced the family history factor [7]. They had shown that 58% of offspring were allergic if both the parents were allergic, while only 20% were allergic if only one parent was allergic, and only 6% of offspring were allergic if none of the parents were allergic. They concluded that heredity appears to contribute significantly in asthma, conferring susceptibility or resistance, affecting the severity or progression of disease, and interacting with environmental influences [8].

Twin studies comparing disease frequency in monozygotic (MZ) versus dizygotic (DZ) twins provide evidence for generic contribution to a disease. Therefore the difference in disease prevalence between MZ and DZ twins would be majorly contributed by genetic factors.

Apart from twin studies, population genetic studies, such as association studies, have played a major role in identification of several causative genes for most of the complex diseases including asthma. Linkage studies, candidate gene studies, and genome-wide association studies (GWAS) have reiterated results on similar line as twin studies, pointing the need to focus on genetics for better understanding of the disease biology of asthma (Fig. 20.1).

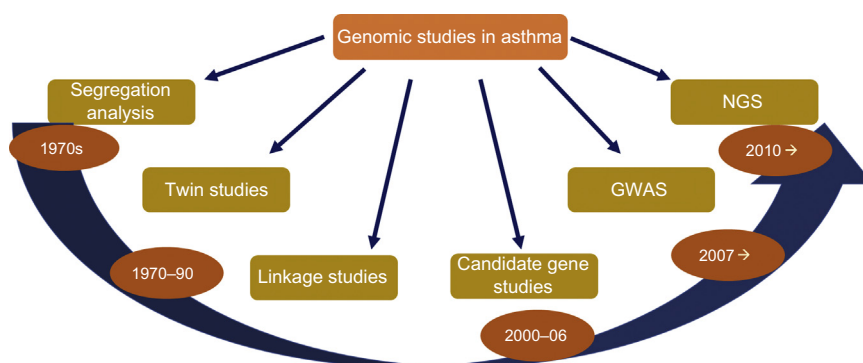
## 20.3 Genetic and genomic studies in bronchial asthma

### 20.3.1 Segregation analysis

To get insights into the genetics of a trait, for example, the number of genes involved and the genetic model: dominant or recessive, polygenic, along with the possibility of external influence of environmental effects, segregation analysis is performed. It helps in estimating the frequency of the trait, its heritability, mode of inheritance and penetrance, indicating the involvement of major genes. Even though segregation analysis is useful in genetic modeling, it has its own limitation in complex diseases. Data from 13,962 families was collected and analyzed by the European Community Respiratory Health Survey Group using complex segregation analysis. They identified that a two-allele gene model with codominant inheritance explained the data best under the assumption that a major gene with low penetrance is involved in the pathogenesis of asthma and therefore found evidence of genetic regulation of asthma [9]. Another study on segregation analysis of 7394 families by Jenkins et al. showed that 15.9% of the individuals had asthma [10]. In 3369 randomly selected individuals from 906 nuclear families, segregation analysis of physician-diagnosed asthma showed evidence of a polygenic or an oligogenic model with evidence of a recessive gene with a high frequency of 0.67 explaining a part of the segregation depending on residual familial effect assessment method [11].

Serum immunoglobulin E (IgE) concentration has been an important component of asthma studies for years. Most of the studies based on segregation analyses of total serum IgE confirmed that their levels were highly heritable. In multiple studies the evidence for recessive inheritance of "high" total serum IgE levels had been found with different estimates for gene frequency and for mean levels in the "low" and "high" phenotypes.

For example, in a northern Dutch population, total elevated serum IgE levels appeared to be inherited in a recessive pattern, on chromosome 5q obtained by sib pair and Logarithm of Odds (LOD) score analysis, in the presence of a two-locus model [12,13]. In a large study of Hispanic families and non-Hispanic white families, Martinez et al. reported evidence for a codominant mode of inheritance in both ethnic groups [14], suggesting a



**FIGURE 20.1** The evolution of genomic research in asthma over the last five decades.



possible way to distinguish carrier of gene from unaffected family members even in the presence of substantial overlap in the distribution of total IgE levels.

For asthma, Martinez and Holberg reported a strong familial polygenic component to asthma, suggesting the presence of several major genes [15]. Following this, multiple studies have reported a strong association between atopy and bronchial hyperresponsiveness [16–18]. There have also been a few segregation analysis studies of other measures of the allergic or asthmatic phenotype. For bronchial responsiveness to methacholine, Townley et al. were unable to exclude genetic or environmental models and concluded that the bimodal distribution of the bronchial response was not the result of segregation at a single autosomal locus [19]. However, in the Dutch population studies, evidence for dominant inheritance was observed after covarying for total serum IgE levels [20,21].

The complexity of the immunological network involved in the pathogenesis of asthma, atopy, and its related traits involve multiple signaling pathways and immunological networks suggesting the involvement of different genes [22]. To identify the key genetic players of asthma and understand their role, different genetic approaches were employed toward identification of these genes, which are discussed in the upcoming sections.

### 20.3.2 Twin genetic studies

Twin studies examine diseases in MZ as well as DZ twins.

MZ twins share the same genetic makeup and a similar environment. DZ twins on the other hand have a 50% similar genetic makeup and identical environmental rearing (if reared together). Twin studies are helpful in studying the heritability of complex genetic diseases having both genetic as well as environmental components. Differences in heritability of MZ versus DZ twins help segregate variance due to environmental factors and genetics, respectively. This is because differences in disease in a DZ twin pair can be due to both genetics and environment whereas in a MZ pair, it is only due to environment. However, epigenetics is an exception. Epigenetics is a variation at transcriptional level, which can skew the heritability results of MZ twin studies. The phenotypic differences and disease discordance in MZ twins at birth may be attributable to epigenetic effects in gene expression. Epigenetic effects are thus studied in context of time (stage in life) of evaluation, pathogenesis phase of disease, and gene locus involved. Most commonly involved epigenetic mechanisms in humans include DNA methylation and histone modification.

There are various twin studies, namely:

- classical twin study that involves studying sets of twins;
- twin adoption study that involves studying twins who have been reared apart (i.e., each of the twins adopted by a different family); and
- twin family study that involves studying the entire family of twin siblings.

Published literatures show that the chances of developing asthma in a cotwin are increased if the index twin has the disease, and this effect is even more pronounced in the cases of MZs than DZs. This depicts the genetic contribution to variance in asthma pathogenesis, with heritability estimation ranging from 48% to 79% according to various published articles [23].

Danish twin registry, established in 1953, is the oldest national twin registry in the world. Thomsen et al. studied 38,782 Danish twins via the twin registry for bronchial asthma. Their key findings included risk factors for developing asthma, which were female sex, atopic dermatitis, hay fever, and smoking—active and passive. Bacillus Calmette Geurin (BCG) vaccination was found to be protective for asthma development. Also, it was observed that higher age of onset of asthma in index twin was seen to be associated with decreased disease risk in the cotwin [24]. In another analysis on the same twins, Body Mass Index (BMI) was also positively correlated to asthma predisposition with obesity being another risk factor [25].

Niemann et al. studied the Finnish twin cohort comprising 27,776 twin individuals. The asthma concordance rate among the MZ twins was found higher for all age-groups starting from 28 years to 60 +. This was one of the few twin registry studies, which confirmed asthma diagnosis by a physician instead of only self-reporting [26].

Puerto Rican Hispanic twins aged 1–3 years were also studied for asthma susceptibility. Bunyavanich et al. found that reducing environmental tobacco smoke exposure for the toddlers can greatly reduce asthma prevalence in them [27].

Another study on Danish twin pairs revealed that the risk of developing asthma in a lower birth weight twin is higher than the cotwin, and this risk increases by 4% with every 100 g decrease in birth weight. This effect of low birth weight on asthma predisposition was present even after adjusting for Apgar score, birth length, and sex [28] (Table 20.1).

**TABLE 20.1** List of twin studies on asthma published over the years.

S. no.	Ethnicity	No. of twin pairs	Gene/mechanism studied	Year	Reference
1.	Finnish	525	Heritability estimates	1991	[26]
2.	Norwegian	5864	Familial risk of asthma	1997	[29]
3.	British	4910	Heritability estimates	2001	[30]
4.	Danish	63	Genetic basis of serum variability of IgE levels in asthma	2001	[31]
5.	Finnish	6012	Variability in hospital admissions among asthmatics	2001	[32]
6.	Dutch	8633	Heritability estimates	2007	[33]
7.	Italian	392	Heritability estimates	2008	[34]
8.	Puerto Ricans	339	Early-life environmental contributions to asthma	2013	[27]
9.	Swedish	5459	In utero environmental contributions to asthma	2015	[35]
10.	Danish	17,391	Age of onset of asthma	2010	[24]

IgE, Immunoglobulin E.

There are a few drawbacks of twin study design, which are as follows:

- Lack of generalizability of results in singletons.
- Most twin studies rely on national twin registries, enrollment in most of them are voluntary. This may lead to recruitment bias.
- Most twin studies are retrospective or cross-sectional. There is a possibility of recall bias.

Further, there are a multitude of published twin studies on asthma [23]; however, we have included few of the important ones in this chapter. Current twin studies are now focusing on epigenetics, epistasis, and metagenomics. These studies will further improve our understanding of asthma genetics. Furthermore, while heritability and environmental influences are recorded by twin studies, linkage studies highlight the chromosomal areas showing asthma association, as discussed in the following segment.

### 20.3.3 Genetic linkage

Genetic linkage is a very powerful tool that has aided identification of genes involved in various diseases. Initially designed to find genes responsible for simple Mendelian diseases, it soon found its application in the identification of genes involved in complex diseases such as asthma. Genetic linkage occurs when two chromosomal loci are physically located close to each other and passed together as a single unit. Linkage analysis examines whether a disease phenotype is inherited jointly with a genetic marker locus, thereby indicating that disease locus is physically located close to marker locus. It relies on using family-based data: family structure information and precisely defined phenotypic data of the families where there is aggregation of the disease and genetic marker data for the individuals in the family. Linkage studies also have two main approaches: candidate gene study or genome-wide linkage. Hypothesis-independent genome-wide approach has been able to identify higher number of novel genes and loci associated to asthma as compared to hypothesis-driven candidate gene approach for linkage analysis in asthma.

Several genome-wide linkage studies have been conducted, and regions on almost every chromosome have been linked to asthma or the associated phenotype. Linkage studies have enabled identification of many asthma susceptibility genes across different populations.

In a study conducted by Ober et al., Hutterites, who belong to a religious isolate of European ancestry, were subjected to genome-wide screening to identify genes that influence asthma and asthma-associated phenotypes [36]. Using 292 autosomal and three X–Y pseudoautosomal markers, genome-wide screening was performed in a primary sample of 361 individuals and a replication sample of 292 individuals. A total of 12 markers in 10 regions were identified, which showed possible linkage to asthma or an associated phenotype using the semi-parametric likelihood ratio,  $\chi^2$  test, and the transmission–disequilibrium test. Markers in four regions (5q23–31, 12q15–24.1, 19q13, and 21q21) with possible linkage in both the primary and replication samples were identified,

which showed linkage to asthma phenotypes in other samples. In addition to this, possible linkage was reported in two adjacent markers in one additional region (3p24.2–22) showing for the first time in the Hutterites.

In another study, polymorphisms in 17q21 were shown to confer higher risk in early onset asthma, which was seen to increase further on exposure to environmental tobacco smoke in early life [37]. After this, four genes in this region were identified with potential role in asthma pathogenesis [38]. Genome-wide significant evidence of linkage to region 5q13 and 6p21 were provided by Teerlink et al. [39].

Multiple groups have identified regions that show evidence for linkage termed replicated regions. These are most commonly found in chromosomes 5q, 6p, 11q, and 12q. Other regions such as 1, 2q, 3, 14, 9, 16, and 17q are replicated in at least two populations.

#### **20.3.3.1 Chromosome 5q**

Multiple genes on chromosome 5q were identified, which may play an important role in pathogenesis of asthma. Initially, linkage to this chromosome was reported by a study based on candidate gene approach, which suggested regulatory aspect of IgE and its inflammatory association with asthma and allergy. Following this, multiple studies on different populations, such as the Dutch [20], Caucasian families [40], and even Hutterites [36], reported linkage of 5q to asthma, Bronchial Hyper-Responsiveness (BHR), and serum IgE levels, signifying its importance. The cytokine gene cluster at 5q contains many cytokines relevant to allergic inflammation, including interleukin (IL)-4, IL-5, and IL-13, which are now validated therapeutic targets. Table 20.2 shows the full list of this cytokine gene cluster that has importantly influenced asthma molecular pathobiology research.

#### **20.3.3.2 Chromosome 6p**

Linkage to asthma and its associated phenotype in chromosome 6 helped in identification of the Major Histocompatibility Complex (MHC) region, especially Human Leukocyte Antigen-G (HLA-G), which was considered to be a major locus in influencing allergic diseases [41]. Apart from this, tumor necrosis factor  $\alpha$  was also identified from such linkage studies [42].

#### **20.3.3.3 Chromosome 11**

Markers on chromosome 11q were detected at the early stages of a genome-wide screen of a broadly defined allergic phenotype indicating evidence of linkage. Sequence variants in the high-affinity IgE receptor (Fc $\epsilon$ R) gene were speculated to increase the risk for development of allergy and possibly asthma [43]. In addition to that, multiple genome-wide screens provided continued evidence for linkage of the atopic phenotype along with asthma phenotype in the United Kingdom and African-American population [40,44].

#### **20.3.3.4 Chromosome 12q**

Candidate gene and genome screen approach provided the first evidence for linkage to 12q for phenotypes ranging from bronchial hyperresponsiveness to allergy [40,45]. Several candidate genes map to this wide region, including IFN- $\gamma$ , nitric oxide synthase, and mast cell growth factor. However, variants in these genes have not been fully evaluated.

Table 20.2 shows the chromosome regions involved in asthma as identified by linkage analysis. More than 100 loci have been linked to asthma [46,47].

Studies from different populations suggest the presence of multiple regions on the genome to contain susceptibility genes of bronchial asthma. However, due to multiple reasons, such as founder effects, admixture, and varying environmental influences, the evidence for any given linkage varies across populations. Nonetheless, it is clear from the reported studies that there are multiple regions in the genome that contain susceptibility genes for asthma and allergy. Dissecting out the genes from the regions suggested by linkage studies is performed using positional cloning approach. This has been discussed in detail in Section 20.3.4.

### **20.3.4 Candidate gene studies**

Asthma is a polygenic disease condition with no single gene as the “asthma-causing gene.” There is an interaction between multiple genes and environmental factors along with a variety of host factors, which lead to the pathophysiology of the disease.

There are studies highlighting the importance of environmental factors on the disease pathology. In the 1990s Martinez et al. studied that the environmental endotoxin levels affected the expression of genetic polymorphism

**TABLE 20.2** List of genes and their chromosomal locations identified by linkage studies.

Chromosomal locations	Genes identified by linkage studies
1p	EGR-1, PTGER-3, CLC-1, VCAM-1, GSTM1
1q	CHIA, FLG, IL-10, A1, TGF- $\beta$ 2
2q	IL-1R1, INPP4A, IL-1RN, CTLA-4, IL-1( $\alpha$ , $\beta$ ), DPP10, IL-8RA
3p	CCR1
4q	IL-8, APA, IL-21
5q	IL-3, 4, 5, 9, 10, 12, 13; CD14; ADRB2; GRP1; CYFIP2; LTC4S
6p	TCR- $\beta$ V, IL-17, HLA-DRB1, TNF- $\alpha$ , ARG1
6q	SOD2, TGF- $\beta$ 1
7p	IL-6, GPRA, TCRG, EGFR
7q	PAI-1, eNOS-NOS3
8p	NAT2
8q	PAF-1
9p	PTPRD
9q	PTGES
10q	PTEN
11p	MUC2
11q	PTGDR, FcR1 $\beta$ , GSTP1, CC16, IL-18
12p	AICDA, CD-69
12q	VDR, STAT6, IRAK3, IL-22, IFNG, nNOS:NOS1,
13q	SETDB2, PHF11, RCC1, CYSLTR2
14q	CMA1, ARG2, PTGER2, AACT
15q	ERK-3
16p	IL-4R
16q	CYBA
17p	ALOX15
17q	iNOS:NOS2, CCL5, ORMDL3, STAT3, CCL11, ACE
18q	SCCA-1
19p	TBXA2R, ICAM-1
19q	PTGER1, TGF $\beta$ 1
20p	ADAM33
20q	CDH26
21q	SOD1, CBR1
22q	GSTT1

in the promoter region of CD14 and thus its clinical asthmatic phenotype. CD14 is a pattern recognition parameter for endotoxins. C260T polymorphism in this gene interacted with house dust endotoxin levels where a protective response was mounted for lower levels for endotoxins, but after a certain threshold of endotoxin levels, the presence of this polymorphism was a risk factor for asthma [48].

Epistasis is the science of studying the phenotypic effect of interaction of two or more genes [49]. Lee et al. studied the interaction of polymorphisms in IL-4 and IL-4 receptor alpha and identified an increased

susceptibility to asthma due to a synergistic interaction between the two genes. A similar interaction is also seen with polymorphisms in IL-13 and IL-4R $\alpha$  [50].

**Positional cloning:** It involves the identification of the position of disease-causing gene on chromosome and its sequence. It takes clues from linkage or association studies regarding the location of gene. Few of asthma-related genes have been discovered via this method, namely, ADAM33, DPP10, CYFIP2, HLAG, GPRA, SFRS8, and PHF11.

ADAM33 is the first asthma candidate gene to be discovered by positional cloning. It belongs to the family of Zn<sup>2+</sup>-dependent matrix metalloproteinases and is expressed in smooth muscle cells of airways. The precise role of ADAM33 in asthma is not confirmed but is thought to be involved in airway remodeling and bronchial hyper-responsiveness. In addition, ADAM33 in early lung development can lead to low lung function [51]. Further, several polymorphisms in the PHF11 gene are associated with high levels of IgE in the blood and thus severe asthma [33]. Another positionally cloned gene, DPP10, is involved in inactivation of voltage-gated potassium channels in the airway epithelium. However, the clear role of DPP10 in asthma pathogenesis has not been elucidated. Unlike other members of the DPP family, protein encoded by DPP10 does not have any detectable protease activity [33].

Candidate genes are genes playing a role in the pathophysiology of a disease. Thus candidate genes for asthma include genes involved in allergen presentation, inflammatory response to allergens, bronchoconstriction, and response to bronchodilators. Usually, candidate genes are identified by sequencing a particular region of DNA, based on an underlying hypothesis related to the function of the region. More than 100 genes have shown to be associated with asthma pathogenesis, thus contributing to disease risk. While this approach is efficient, it reversed the previous trends where new molecular pathobiology of asthma was discovered through genomic insights. Here, experimental pathobiological insights, often from mouse models of disease that partly recapitulate human asthma, drove genetic understanding. Intersections between positive hits from candidate genes and positional linkages were most robust. Some of the genes, which have been replicated by various researchers and have shown to be associated with asthma disease pathology, are listed in Table 20.3.

ADRB2 is a beta 2 adrenoreceptor gene. Many researchers have recognized an association between polymorphisms in ADRB2 and bronchodilator response to short-acting and long-acting ADRB2 agonists. However, many researchers have not been able to replicate these results. Clear mechanistic understanding of role of ADRB2 on bronchodilator response can help in personalizing therapy for patients [52].

Even though numerous candidate genes for asthma have been discovered and studied, the contribution of each gene to overall disease pathology is small. There are epidemiological differences in prevalence of different variants among the population. Apart from this, epistatic and epigenetic interactions also add to the overall phenotypic variability. Use of candidate gene knowledge to guide pharmacological basis of therapy is gaining popularity [53,54]. More elaborate research is required to take asthma genetics to the clinic for personalized therapy. Apart from selecting candidate genes on the basis of existing hypothesis, a genome-wide approach to identify nonintuitive yet important locations for asthma pathogenesis in the genome could be achieved by GWAS, as explained in the following section.

### 20.3.5 Genome-wide association studies

GWAS works on the concept of “common variant–common disease” wherein it is hypothesized that common variants/polymorphisms with small effect sizes can identify regions associated with a common disease in a given population. Using predesigned or custom-made panel of a million single-nucleotide polymorphisms (SNPs), GWAS exploits the chances of identifying association of SNPs to a given disease in comparison to normal subjects.

Over the past two decades, GWAS has enabled identification of hundreds of common risk alleles for complex human diseases. This approach has enabled critical information in many complex diseases, for example, highlighting the role of complement genes in age-related macular degeneration, Crohn’s disease, or of regulatory proteins in blood lipid levels. The only drawback is lack of translation of such findings as most of these common variants have subtle functional consequences.

In 2007 the first GWAS on asthma childhood cohort was performed and published by Moffatt et al. in Nature [55]. This was the first study that helped in the identification of a novel loci strongly associated with asthma biology ( $P < 10^{-12}$ ), that is 17q21, wherein one of the most promising and well-replicated childhood asthma gene ORMDL3–GSDMB could be identified. Since then many GWAS have been published in the context of asthma



**TABLE 20.3** Candidate genes and their chromosomal location.

Chromosomal location	Candidate gene	Name of gene
1p3	ATPAF1	Mitochondrial protein binding to beta subunit of F1-ATP synthase
1q	CHI3L1	Chitinase-3-like 1
2q	CD-28	Cluster of differentiation-28
2q	DPP10	Dipeptidyl peptidase
3p	Chemokine cluster	
5q	GRL	Glucocorticoid receptor
5q	CYFIP2	Cytoplasmic-FMR1-interacting protein 2
5q	IL-3–5, -9, -10, -13	Interleukin 3–5, 9, 10, 13
5q31	CD14	Cluster of differentiation-14
5q	SPINK5	Serine peptidase inhibitor, Kazal type 5
5q	TIM1	T-cell immunoglobulin and mucin domain 1
5q31	ADRB2	Adrenoreceptor beta 2
6p21–p22	TNF	Tumor necrosis factor
6p21–p22	HLA region–HLA	Human leukocyte antigen region
6p	LTA	Lymphotoxin-alpha
7p	NOD1	Nucleotide-binding oligomerization domain containing 1
7p	GPRA	G-protein-coupled receptor for asthma
11q	FCER1B	High-affinity IgE receptor
11q	SGB1A1	Secretoglobulin family 1A member 1
12q23	NOS	Nitric oxide synthase
12q23	IFN gamma	Interferon gamma
12q	STAT6	Signal transducer and activator of transcription 6
13	PFH11	PHD finger protein 11
16p	IL-4RA	IL-4 receptor
17q	CCL5	C–C motif chemokine ligand 5
17q	ORMDL3	ORMDL sphingolipid biosynthesis regulator 1
20p	ADAM33	A disintegrin and metalloproteinase 33

and allergy, with >24 studies performed on asthma alone [56–58]. Details of GWAS on asthma, key findings of which were replicated in other studies, have been tabulated in Table 20.4.

Although GWAS helped in identifying some key novel genes involved in diseases, most of the reported variants when replicated in separate cohorts failed to reach significance. The key gain from GWAS over previous candidate gene studies was that the unbiased nature of GWAS led to new pathobiological insights, including discovery of orphan genes of no known function at the time such as ORMDL3. Some of the most successful leads in asthma, for example, the cytokine thymic stromal lymphopoietin (TSLP), which is now an exciting therapeutic target, came from an intersection of GWAS-based genetic significance [61], along with a plethora of preceding molecular evidence that it is secreted from barrier epithelial surfaces and orchestrates immunological inflammation. This and other GWAS findings led to an important shift from inflammatory cells to airway epithelium as the key orchestrator of asthma pathobiology, an insight that explained many gaps in previous models, which failed to explain the differences between atopy and asthma.

**TABLE 20.4** Findings of asthma genome-wide association studies that were replicated in other studies.

Chromosome	Gene name	SNP	P-value	Year	Reference
2	IL-18R1	rs3771166	3.4E – 09	2010	[59]
4	TLR1	rs4833095	5.0E – 12	2013	[60]
5	TSLP	rs1837253	7.3E – 10	2011	[61]
5	RAD50	rs6871536	2.4E – 09	2011	[62]
6	HLA-DQA1–HLA-DQB1	rs17843604	1.7E – 10	2010	[59]
7	CDHR3	rs6967330	1.4E – 08	2013	[63]
8	MIR5708–ZBTB10	rs7009110	4.0E – 09	2013	[60]
9	IL-33	rs1342326	9.2E – 10	2010	[59]
11	WNT11–LRRC32	rs7130588	1.8E – 08	2011	[62]
15	RORA	rs11071559	3.8E – 09	2011	[62]
15	SMAD3	rs744910	3.9E – 09	2010	[59]
16	CLEC16A	rs62026376	1.0E – 08	2013	[60]
17	GSDMB	rs7216389	9.0E – 11	2007	[55]

SNP, Single-nucleotide polymorphism.

While here we touch upon the success stories, there have been many failures of GWAS hits as well, especially in other diseases. There are a number of reasons for the failure of GWAS in general and their limited success in the context of asthma:

1. The SNPs used for panels were initially designed for Caucasian populations but have been since used for all other populations. Since minor allele frequencies vary with each population, it is highly unlikely for the same set of SNPs to be relevant across all populations.
2. Smaller sample size leads to lower statistical power to detect significant associations and higher probability of false positives.
3. Also, GWAS primarily detects common risk variants and therefore loses out on information of more deleterious rare variants, which are likely to be causal variants in any complex disease.
4. Replication cohorts are mostly built with subjects from different ethnic backgrounds. This change in population substructure and exposure to different environments could possibly affect penetration and expressivity of variants leading to poor replication.
5. Also, variability in endotypes of case subjects along with variable study designs further leads to decrease in probability of gaining consensus.

Success of GWAS strongly depends on the number of subjects employed in the study to achieve high statistical confidence, which is also called the power of the study. Risk allele frequency and odds ratio can help determine the proportion of variance in disease liability, which gives an estimate regarding the strength of association, that is, power. Disease heterogeneity and misclassification can affect power calculation, for example, subjects having different subtypes of asthma may be classified together along with controls who do not have asthma but other allergic diseases. Such studies can significantly decrease the power of study for asthma by including genetic variants that are also important in other allergic diseases. In studies where cases and controls are clearly segregated, there is a higher likelihood of finding true associations, for example, identification of ORMDL3 in childhood asthma and CDHR3 in children with early onset and severe asthma [63–65].

Also, in complex diseases wherein gene–environment interaction has a significant role to play, the loss of such relevant information could lead to underestimation of the SNP effect. Studies with children having rhinovirus, wheezing illness, or early-life tobacco exposure showed genetic variants in ORMDL3, which were not replicated in large independent studies.

For asthma, fewer genome-wide associations have been identified with confidence as compared to other complex diseases such as Type II diabetes and schizophrenia [57]. Relatively high disease prevalence and insufficient

sample size leading to low statistical power, along with limitation in study design, missing out on asthma subtypes and noninclusion of information regarding gene–environment interactions, could be the possible factors. Due to these issues and decrease in per sample sequencing cost, the focus on deciphering asthma genomics eventually shifted to exome- and genome-sequencing strategies for identifying asthma genetic variants.

### 20.3.6 Next-generation sequencing

Next-generation sequencing (NGS) is a low-cost platform for sequencing DNA using simultaneous parallel sequencing reactions for short DNA fragments. It is possible to perform whole-genome, whole-exome, and targeted-region sequencing using the NGS technology. Different companies utilize different techniques for sequencing these short fragments [66,67]. The companies currently marketing next-generation automated techniques are as follows:

- ABI SOLiD platform: sequencing by ligation;
- DNA nanoball sequencing by BGI Retrovoluty: sequencing by hybridization and ligation;
- Illumina/Solexa Genome Analyzer: sequencing by synthesis;
- Ion Torrent: Proton/PGM sequencing: sequencing by synthesis; and
- Roche GS-FLX 454 Genome Sequencer: pyrosequencing.

The details of each of the abovementioned platforms are beyond the scope of this chapter.

A few studies have been conducted utilizing high-throughput NGS for asthma risk assessment, diagnosis, and treatment guidance.

NGS of whole exome was performed on all members (total six) of a family segregating asthma. DeWan et al. identified 10 novel, nonsynonymous variants that segregated perfectly with asthma. Most of these variants were located in three asthma candidate genes, that is, PDE4DIP, CBLB, and KALRN. In addition, a trend of more rare variants in asthma candidate genes among case children than control children was recorded. However, the commonly known risk variants (such as rs7216389 in ORMDL3) did not segregate between the asthma-affected and nonaffected individuals. This was the first whole-exome sequencing study performed for variant identification in an asthmatic family [68].

Dannemiller et al. conducted a prospective nested case–control study to analyze the house dust, moisture, and mold present in a household and relate it to later life asthma risk. The dust was used to sequence the DNA present and analyze fungal concentration and diversity. It was observed that lower fungal diversity of different taxonomic groups leads to an increased risk of asthma development in later life probably due to immunological priming. Lower fungal diversity of genus *Cryptococcus* showed the most significant effect. The DNA sequencing was performed using next-generation titanium pyrosequencing platform [69].

Losada et al. studied the nasal microenvironment using NGS to characterize the transient and resident bacteria involved in asthma pathogenesis. *Moraxella*, *Staphylococcus*, *Streptococcus*, *Hemophilus*, and *Fusobacterium* accounted for the major nasal microbial signature of the individuals. Interseasonal differences were also observed. However, further studies are required to clinically correlate their results [70].

Leung et al. studied asthma-related haplotypes using next-generation 454 pyrosequencing technology in different ethnic groups namely European, Chinese, and Caucasians. Variations in alleles and polymorphism frequency were observed in different ethnic groups, due to epidemiological differences in distribution of asthma genes thereby pointing toward evolution in population genetics [71].

Further, Burnstein et al. studied diisocyanate-induced occupational asthma using the NGS platform. They discovered 130 functional noncoding SNPs in previously identified gene loci (with respect to diisocyanate-induced asthma). Five of these SNPs (rs11571537, rs2446824, rs2287231, rs2513789, and rs147978008) were found to have a potential regulatory role as confirmed by electrophoretic mobility shift assay as well as luciferase reporter activity [72].

Campbell et al. performed whole-genome sequencing on DNA of family-enriched asthmatic individuals and discovered a 6 kb pair intron deletion in gene NEDD4L to be associated with the disease phenotype. However, this deletion is quite rare with a frequency of only 0.6% in the Hutterite family studied. A gene-knockout experiment on mice for NEDD4L gene exhibited mucus hypersecretion and inflammation. The study had a limitation of small sample size, and replication studies are required for studying such rare variants [73].

Sheu et al. studied the airway epithelial homeostasis using NGS and genome expression microarray and discovered myocyte enhancer factor 2C (MEF2C), MAM domain containing glycosylphosphatidylinositol

anchor 1 (MDGA1), and  $K^+$  voltage-gated channel subfamily J member 2 (KCNJ2) genes to have a significant role. MEF2C is expressed in bronchial epithelial cells and is involved in mitogen-activated protein kinase pathway leading to decreased cellular proliferation. MEF2C is downregulated in asthmatic patients. MDGA1 is repressed in asthmatic individuals via a microRNA and thus results in suppressed cell–cell adhesion in bronchial epithelium. KCNJ2 leads to bronchial epithelial cell apoptosis in asthmatic individuals. Expression profiles of microRNAs and messenger RNAs and their interaction with the mentioned genes provides a new perspective in studying asthma pathophysiology [74].

Evidently, the current role of NGS in asthma pathophysiology is limited to the following:

- discovery of new variants—rare variants as a part of missing heritability;
- understanding the microbiological environment of the pulmonary system as well as the external environment and its interaction with the disease; and
- RNA-expression profiling and their interaction with candidate genes to understand pathophysiology of asthma better.

## 20.4 Management and treatment of bronchial asthma

Comprehensive management and treatment of bronchial asthma requires coordinated approach of primary- and secondary-care clinicians and respiratory health-care professionals. The main strategy includes symptomatic control supported by adequate inflammation control. Role of antibiotics is limited to specific antimicrobial therapy.

### 20.4.1 Symptomatic control

Bronchodilators help in symptomatic control of disease by reversing bronchoconstriction. Short-acting beta 2 agonists, such as albuterol, provide immediate relief and are thereby used as rescue medicine. Long-acting beta 2 agonists, such as formoterol, are prescribed to provide better control for regular use.

### 20.4.2 Inflammation control

Corticosteroids and leukotriene receptor antagonists help in controlling inflammation. Hence, they do not provide immediate relief but help in controlling future episodes. If substantial relief is not produced and exacerbations continue, then drugs are added subsequently. Step 1 starts with inhaled corticosteroids and subsequently oral corticosteroids followed by immunological target therapy in selected patients.

### 20.4.3 Commonly used drugs in bronchial asthma

Drugs used to control asthma include:

- beta 2 adrenergic receptor agonists (bronchodilator);
- anticholinergics (bronchodilator);
- phosphodiesterase inhibitor (bronchodilators);
- inhaled corticosteroids;
- oral corticosteroids;
- leukotriene receptor antagonist—montelukast; and
- monoclonal antibodies against immunological molecular targets such as IgE, IL-4, IL-13, IL-5, and TSLP.

## 20.5 Conclusion

So far, numerous studies have suggested the importance of genetics in bronchial asthma. Starting from twin studies to utilizing NGS approaches, a massive effort has been taken to understand asthma pathophysiology. Many of the genes discovered from these genomic approaches are now well-established therapeutic targets. The journey from a genomic hit to a validated asthma gene and potential therapeutic target is beyond the scope of

this chapter, but the road is littered with false discoveries, unmet expectations, and surely missed opportunities. It seems certain that the future of complex disease genomics will be incomplete without integrating the contribution of the environment. A system biology approach that combines information from multiple “omics” studies (genomics, transcriptomics, epigenetics, metabolomics, etc.) will be crucial for advancing the field further.

## References

- [1] Reddel HK, Bateman ED, Becker A, Boulet LP, Cruz AA, Drazen JM, et al. A summary of the new GINA strategy: a roadmap to asthma control. *Eur Respir J* 2015;46(3):622–39. Available from: <https://doi.org/10.1183/13993003.00853-2015>.
- [2] Loftus PA, Wise SK. Epidemiology of asthma. *Curr Opin Otolaryngol Head Neck Surg* 2016;24(3):245–9. Available from: <https://doi.org/10.1097/MOO.0000000000000262>.
- [3] Bijanzadeh M, Mahesh PA, Ramachandra NB. An understanding of the genetic basis of asthma. *Indian J Med Res* 2011;134:149–61.
- [4] Sandford A, Weir T, Pare P. The genetics of asthma. *Am J Respir Crit Care Med* 1996;153(6 Pt 1):1749–65. Available from: <https://doi.org/10.1164/ajrccm.153.6.8665031>.
- [5] Doull IJ, Lawrence S, Watson M, Begishvili T, Beasley RW, Lampe F, et al. Allelic association of gene markers on chromosomes 5q and 11q with atopy and bronchial hyperresponsiveness. *Am J Respir Crit Care Med* 1996;153(4 Pt 1):1280–4. Available from: <https://doi.org/10.1164/ajrccm.153.4.8616554>.
- [6] Cooke, Robert A, Albert Vander Veer. Human Sensitization. *The Journal of Immunology* 1916;1(3):201–305.
- [7] Van Arsdel Jr. PP, Motulsky AG. Frequency and heritability of asthma and allergic rhinitis in college students. *Acta Genet Stat Med* 1959;9:101–14.
- [8] Zaas D, Schwartz DA. Genetics of environmental asthma. *Semin Respir Crit Care Med* 2003;24(2):185–96. Available from: <https://doi.org/10.1055/s-2003-39017>.
- [9] [No authors listed]. Genes for asthma? An analysis of the European Community Respiratory Health Survey. *Am J Respir Crit Care Med* 1997;156(6):1773–80. <<https://doi.org/10.1164/ajrccm.156.6.9611068>>.
- [10] Jenkins MA, Hopper JL, Giles GG. Regressive logistic modeling of familial aggregation for asthma in 7,394 population-based nuclear families. *Genet Epidemiol* 1997;14(3):317–32 <[https://doi.org/10.1002/\(SICI\)1098-2272\(1997\)14:3<317::AID-GEPI9>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1098-2272(1997)14:3<317::AID-GEPI9>3.0.CO;2-1)>.
- [11] Holberg CJ, Elston RC, Halonen M, Wright AL, Taussig LM, Morgan WJ, et al. Segregation analysis of physician-diagnosed asthma in Hispanic and non-Hispanic white families. A recessive component? *Am J Respir Crit Care Med* 1996;154(1):144–50. Available from: <https://doi.org/10.1164/ajrccm.154.1.8680670>.
- [12] Meyers DA, Postma DS, Panhuysen CI, Xu J, Amelung PJ, Levitt RC, et al. Evidence for a locus regulating total serum IgE levels mapping to chromosome 5. *Genomics* 1994;23(2):464–70. Available from: <https://doi.org/10.1006/geno.1994.1524>.
- [13] Xu J, Levitt RC, Panhuysen CI, Postma DS, Taylor EW, Amelung PJ, et al. Evidence for two unlinked loci regulating total serum IgE levels. *Am J Hum Genet* 1995;57(2):425–30.
- [14] Martinez FD, Holberg CJ, Halonen M, Morgan WJ, Wright AL, Taussig LM. Evidence for Mendelian inheritance of serum IgE levels in Hispanic and non-Hispanic white families. *Am J Hum Genet* 1994;55(3):555–65.
- [15] Martinez FD, Holberg CJ. Segregation analysis of physician-diagnosed asthma in Hispanic and non-Hispanic white families. *Clin Exp Allergy* 1995;25(Suppl. 2):68–70 Discussion 95–96.
- [16] Collins SA, Pike KC, Inskip HM, Godfrey KM, Roberts G, Holloway JW, et al. Validation of novel wheeze phenotypes using longitudinal airway function and atopic sensitization data in the first 6 years of life: evidence from the Southampton Women’s survey. *Pediatr Pulmonol* 2013;48(7):683–92. Available from: <https://doi.org/10.1002/ppul.22766>.
- [17] Mohan A, Roberto AJ, Whitehill BC, Mohan A, Kumar A. Exploration of the genetic epidemiology of asthma: a review, with a focus on prevalence in children and adolescents in the Caribbean. *West Indian Med J* 2014;63(7):687–92. Available from: <https://doi.org/10.7727/wimj.2014.115>.
- [18] Sheikh A. Itch, sneeze and wheeze: the genetics of atopic allergy. *J R Soc Med* 2002;95(1):14–17.
- [19] Townley RG, Bewtra A, Wilson AF, Hopp RJ, Elston RC, Nair N, et al. Segregation analysis of bronchial response to methacholine inhalation challenge in families with and without asthma. *J Allergy Clin Immunol* 1986;77(1 Pt 1):101–7.
- [20] Postma DS, Bleeker ER, Amelung PJ, Holroyd KJ, Xu J, Panhuysen CI, et al. Genetic susceptibility to asthma—bronchial hyperresponsiveness coinherited with a major gene for atopy. *N Engl J Med* 1995;333(14):894–900. Available from: <https://doi.org/10.1056/NEJM199510053331402>.
- [21] Wiesch DG, Meyers DA, Bleeker ER. Genetics of asthma. *J Allergy Clin Immunol* 1999;104(5):895–901.
- [22] Los H, Koppelman GH, Postma DS. The importance of genetic influences in asthma. *Eur Respir J* 1999;14(5):1210–27.
- [23] Thomsen SF. Exploring the origins of asthma: lessons from twin studies. *Eur Clin Respir J* 2014;1(Suppl. 1). Available from: <https://doi.org/10.3402/ecrj.v1.25535>.
- [24] Thomsen SF, Duffy DL, Kyvik KO, Backer V. Genetic influence on the age at onset of asthma: a twin study. *J Allergy Clin Immunol* 2010;126(3):626–30. Available from: <https://doi.org/10.1016/j.jaci.2010.06.017>.
- [25] Thomsen SF, Duffy DL, Kyvik KO, Skytthe A, Backer V. Risk of asthma in adult twins with type 2 diabetes and increased body mass index. *Allergy* 2011;66(4):562–8. Available from: <https://doi.org/10.1111/j.1398-9995.2010.02504.x>.
- [26] Nieminen MM, Kaprio J, Koskenvuo M. A population-based study of bronchial asthma in adult twin pairs. *Chest* 1991;100(1):70–5. Available from: <https://doi.org/10.1378/chest.100.1.70>.
- [27] Bunyavanich S, Silberg JL, Lasky-Su J, Gillespie NA, Lange NE, Canino G, et al. A twin study of early-childhood asthma in Puerto Ricans. *PLoS One* 2013;8(7):e68473. Available from: <https://doi.org/10.1371/journal.pone.0068473>.
- [28] Kindlund K, Thomsen SF, Stensballe LG, Skytthe A, Kyvik KO, Backer V, et al. Birth weight and risk of asthma in 3-9-year-old twins: exploring the fetal origins hypothesis. *Thorax* 2010;65(2):146–9. Available from: <https://doi.org/10.1136/thx.2009.117101>.



- [29] Harris JR, Magnus P, Samuelsen SO, Tambs K. No evidence for effects of family environment on asthma. A retrospective study of Norwegian twins. *Am J Respir Crit Care Med* 1997;156(1):43–9. Available from: <https://doi.org/10.1164/ajrccm.156.1.9609094>.
- [30] Koeppen-Schomerus G, Stevenson J, Plomin R. Genes and environment in asthma: a study of 4 year old twins. *Arch Dis Child* 2001;85(5):398–400. Available from: <https://doi.org/10.1136/adc.85.5.398>.
- [31] Jacobsen HP, Herskind AM, Nielsen BW, Husby S. IgE in unselected like-sexed monozygotic and dizygotic twins at birth and at 6 to 9 years of age: high but dissimilar genetic influence on IgE levels. *J Allergy Clin Immunol* 2001;107(4):659–63. Available from: <https://doi.org/10.1067/mai.2001.113565>.
- [32] Huovinen E, Kaprio J. Twins and asthma. Difference in admission rates may be due to other factors. *Brit Med J* 2001;322(7285):556.
- [33] van Beijsterveldt CE, Boomsma DI. Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins. *Eur Respir J* 2007;29(3):516–21. Available from: <https://doi.org/10.1183/09031936.00065706>.
- [34] Fagnani C, Annesi-Maesano I, Brescianini S, D'Ippolito C, Medda E, Nistico L, et al. Heritability and shared genetic effects of asthma and hay fever: an Italian study of young twins. *Twin Res Hum Genet* 2008;11(2):121–31. Available from: <https://doi.org/10.1375/twin.11.2.121>.
- [35] Orqvist AK, Ullemar V, Lundholm C, Kuja-Halkola R, Magnusson PKE, Lichtenstein P, et al. Fetal growth and childhood lung function in the Swedish twin study on prediction and prevention of asthma. *Ann Am Thorac Soc* 2017;14(7):1147–53. Available from: <https://doi.org/10.1513/AnnalsATS.201611-908OC>.
- [36] Ober C, Cox NJ, Abney M, Di Rienzo A, Lander ES, Changyaleket B, et al. Genome-wide search for asthma susceptibility loci in a founder population. The collaborative study on the genetics of asthma. *Hum Mol Genet* 1998;7(9):1393–8. Available from: <https://doi.org/10.1093/hmg/7.9.1393>.
- [37] Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. *N Engl J Med* 2008;359(19):1985–94. Available from: <https://doi.org/10.1056/NEJMoa0806604>.
- [38] Kumar A, Ghosh B. Genetics of asthma: a molecular biologist perspective. *Clin Mol Allergy* 2009;7:7. Available from: <https://doi.org/10.1186/1476-7961-7-7>.
- [39] Teerlink CC, Camp NJ, Bansal A, Crapo R, Hughes D, Kort E, et al. Significant evidence for linkage to chromosome 5q13 in a genome-wide scan for asthma in an extended pedigree resource. *Eur J Hum Genet* 2009;17(5):636–43. Available from: <https://doi.org/10.1038/ejhg.2008.236>.
- [40] [No authors listed]. A genome-wide search for asthma susceptibility loci in ethnically diverse populations. The collaborative study on the genetics of asthma (CSGA). *Nat Genet* 1997;15(4):389–92. <<https://doi.org/10.1038/ng0497-389>>.
- [41] Kontakioti E, Domvri K, Papakosta D, Daniilidis M. HLA and asthma phenotypes/endotypes: a review. *Hum Immunol* 2014;75(8):930–9. Available from: <https://doi.org/10.1016/j.humimm.2014.06.022>.
- [42] Shin HD, Park BL, Kim LH, Jung JH, Wang HJ, Kim YJ, et al. Association of tumor necrosis factor polymorphisms with asthma and serum total IgE. *Hum Mol Genet* 2004;13(4):397–403. Available from: <https://doi.org/10.1093/hmg/ddh036>.
- [43] Kim ES, Kim SH, Kim KW, Park HS, Shin ES, Lee JE, et al. Involvement of Fc(epsilon)RIbeta gene polymorphisms in susceptibility to atopy in Korean children with asthma. *Eur J Pediatr* 2009;168(12):1483–90. Available from: <https://doi.org/10.1007/s00431-009-0960-x>.
- [44] Daniels SE, Bhattacharaya S, James A, Leaves NI, Young A, Hill MR, et al. A genome-wide search for quantitative trait loci underlying asthma. *Nature* 1996;383(6597):247–50. Available from: <https://doi.org/10.1038/383247a0>.
- [45] Barnes KC, Neely JD, Duffy DL, Freidhoff LR, Breazeale DR, Schou C, et al. Linkage of asthma and total serum IgE concentration to markers on chromosome 12q: evidence from Afro-Caribbean and Caucasian populations. *Genomics* 1996;37(1):41–50. Available from: <https://doi.org/10.1006/geno.1996.0518>.
- [46] Haagerup A, Bjerke T, Schiotz PO, Binderup HG, Dahl R, Kruse TA. Asthma and atopy—a total genome scan for susceptibility genes. *Allergy* 2002;57(8):680–6.
- [47] Malerba G, Pignatti PF. A review of asthma genetics: gene expression studies and recent candidates. *J Appl Genet* 2005;46(1):93–104.
- [48] Zambelli-Weiner A, Ehrlich E, Stockton ML, Grant AV, Zhang S, Levett PN, et al. Evaluation of the CD14/-260 polymorphism and house dust endotoxin exposure in the Barbados Asthma Genetics Study. *J Allergy Clin Immunol* 2005;115(6):1203–9. Available from: <https://doi.org/10.1016/j.jaci.2005.03.001>.
- [49] Yoshikawa T, Kanazawa H, Fujimoto S, Hirata K. Epistatic effects of multiple receptor genes on pathophysiology of asthma—its limits and potential for clinical application. *Med Sci Monit* 2014;20:64–71. Available from: <https://doi.org/10.12659/MSM.889754>.
- [50] Lee SG, Kim BS, Kim JH, Lee SY, Choi SO, Shim JY, et al. Gene-gene interaction between interleukin-4 and interleukin-4 receptor alpha in Korean children with asthma. *Clin Exp Allergy* 2004;34(8):1202–8. Available from: <https://doi.org/10.1111/j.1365-2222.2004.02015.x>.
- [51] Van Eerdedewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002;418(6896):426–30. Available from: <https://doi.org/10.1038/nature00878>.
- [52] Liang SQ, Chen XL, Deng JM, Wei X, Gong C, Chen ZR, et al. Beta-2 adrenergic receptor (ADRB2) gene polymorphisms and the risk of asthma: a meta-analysis of case-control studies. *PLoS One* 2014;9(8):e104488. Available from: <https://doi.org/10.1371/journal.pone.0104488>.
- [53] Gupta A, Gupta LK, Rehan HS, Prakash A. Response to inhaled corticosteroids on serum CD28, quality of life, and peak expiratory flow rate in bronchial asthma. *Allergy Asthma Proc* 2017;38(2):13–18. Available from: <https://doi.org/10.2500/aap.2017.38.4022>.
- [54] Lima JJ, Blake KV, Tantisira KG, Weiss ST. Pharmacogenetics of asthma. *Curr Opin Pulm Med* 2009;15(1):57–62. Available from: <https://doi.org/10.1097/MCP.0b013e32831da8be>.
- [55] Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;448(7152):470–3. Available from: <https://doi.org/10.1038/nature06014>.
- [56] Kim KW, Ober C. Lessons learned from GWAS of asthma. *Allergy Asthma Immunol Res* 2019;11(2):170–87. Available from: <https://doi.org/10.4168/aaair.2019.11.2.170>.
- [57] Vicente CT, Revez JA, Ferreira MAR. Lessons from ten years of genome-wide association studies of asthma. *Clin Transl Immunol* 2017;6(12):e165. Available from: <https://doi.org/10.1038/cti.2017.54>.

- [58] Yan Q, Brehm J, Pino-Yanes M, Forno E, Lin J, Oh SS, et al. A meta-analysis of genome-wide association studies of asthma in Puerto Ricans. *Eur Respir J* 2017;49(5). Available from: <https://doi.org/10.1183/13993003.01505-2016>.
- [59] Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genome-wide association study of asthma. *N Engl J Med* 2010;363(13):1211–21. Available from: <https://doi.org/10.1056/NEJMoa0906312>.
- [60] Ferreira MA, Matheson MC, Tang CS, Granell R, Ang W, Hui J, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J Allergy Clin Immunol* 2014;133(6):1564–71. Available from: <https://doi.org/10.1016/j.jaci.2013.10.030>.
- [61] Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;43(9):887–92. Available from: <https://doi.org/10.1038/ng.888>.
- [62] Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011;378(9795):1006–14. Available from: [https://doi.org/10.1016/S0140-6736\(11\)60874-X](https://doi.org/10.1016/S0140-6736(11)60874-X).
- [63] Bonnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet* 2014;46(1):51–5. Available from: <https://doi.org/10.1038/ng.2830>.
- [64] Ono JG, Worgall TS, Worgall S. 17q21 locus and ORMDL3: an increased risk for childhood asthma. *Pediatr Res* 2014;75(1–2):165–70. Available from: <https://doi.org/10.1038/pr.2013.186>.
- [65] Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med* 2019;7(1):20–34. Available from: [https://doi.org/10.1016/S2213-2600\(18\)30389-8](https://doi.org/10.1016/S2213-2600(18)30389-8).
- [66] Mardis ER. A decade's perspective on DNA sequencing technology. *Nature* 2011;470(7333):198–203. Available from: <https://doi.org/10.1038/nature09796>.
- [67] Metzker ML. Sequencing technologies—the next generation. *Nat Rev Genet* 2010;11(1):31–46. Available from: <https://doi.org/10.1038/nrg2626>.
- [68] DeWan AT, Egan KB, Hellenbrand K, Sorrentino K, Pizzoferrato N, Walsh KM, et al. Whole-exome sequencing of a pedigree segregating asthma. *BMC Med Genet* 2012;13:95. Available from: <https://doi.org/10.1186/1471-2350-13-95>.
- [69] Dannemiller KC, Mendell MJ, Macher JM, Kumagai K, Bradman A, Holland N, et al. Next-generation DNA sequencing reveals that low fungal diversity in house dust is associated with childhood asthma development. *Indoor Air* 2014;24(3):236–47.
- [70] Perez-Losada M, Alamri L, Crandall KA, Freishtat RJ. Nasopharyngeal microbiome diversity changes over time in children with asthma. *PLoS One* 2017;12(1):e0170543. Available from: <https://doi.org/10.1371/journal.pone.0170543>.
- [71] Leung TF, Ko FW, Sy HY, Tsui SK, Wong GW. Differences in asthma genetics between Chinese and other populations. *J Allergy Clin Immunol* 2014;133(1):42–8. Available from: <https://doi.org/10.1016/j.jaci.2013.09.018>.
- [72] Bernstein DI, Lummus ZL, Kesavalu B, Yao J, Kottyan L, Miller D, et al. Genetic variants with gene regulatory effects are associated with diisocyanate-induced asthma. *J Allergy Clin Immunol* 2018;142(3):959–69. Available from: <https://doi.org/10.1016/j.jaci.2018.06.022>.
- [73] Campbell CD, Mohajeri K, Malig M, Hormozdiari F, Nelson B, Du G, et al. Whole-genome sequencing of individuals from a founder population identifies candidate genes for asthma. *PLoS One* 2014;9(8):e104396. Available from: <https://doi.org/10.1371/journal.pone.0104396>.
- [74] Sheu CC, Tsai MJ, Chen FW, Chang KF, Chang WA, Chong IW, et al. Identification of novel genetic regulations associated with airway epithelial homeostasis using next-generation sequencing data and bioinformatics approaches. *Oncotarget* 2017;8(47):82674–88. Available from: <https://doi.org/10.18632/oncotarget.19752>.

# Molecular systems in inflammatory bowel disease

*Saad Pathan*

University of Cambridge Metabolic Research Laboratories/NIHR Cambridge Biomedical Research Centre, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, United Kingdom

## 21.1 Introduction

Ulcerative colitis (UC) and Crohn's disease (CD) are chronic inflammatory diseases of the intestine, which can begin at any age, even in children. Both diseases are commonly known as inflammatory bowel diseases (IBDs). They represent a major challenge, because their causation is unclear, and though there have been significant advances in medical treatment, it is still far from satisfactory, and many patients require major surgery. The clinical phenotype of the patients that respond to treatment compared to those that do not respond can be very similar prior to treatment, and therefore the need is for biomarkers to differentiate between these groups. The genomic portrait of IBD was built with the high-resolution detection of loci that individually have a very modest influence. There is evidence for over 200 distinct genetic loci; however, they are not associated with prognosis. More specifically, there are at least 30 distinct loci that are associated with the phenotype of CD, and 23 loci with UC. Moreover, the boundaries between CD and UC are blurred by the larger proportion of loci that are shared. There are also parallels and paradoxes vis-à-vis other diseases, such as inflammatory, autoimmune, infectious, and primary immunodeficiencies. The IBD-associated variants within the genes, such as *NOD2*, *IL23R*, *ATG16L1*, and *CARD9*, have the potential to adversely affect innate and adaptive immunity, as well as mucosal barrier integrity and autophagy. The functional insights have largely been derived from reductive approaches, such as cell culture and animal models. Unusually severe perturbations of the biology have also been observed in exceedingly rare cases of the monogenic forms of the disease. The molecular systems approach is poised to transform our holistic perception of the polygenic interactions by integrating the multiomic approaches rather than deconstruct the biology albeit the one-dimensional view. New methodologies for mathematical modeling and the exponential increase in computational power have made the integrative approaches of molecular systems biology feasible. The unprecedented scale of the massive data combined with the discovery of new IBD-associated biomarkers is envisioned for prediction, prognosis, and the diagnosis of individuals that are susceptible. The new insights from unraveling the multitude of disease mechanisms are prerequisite for novel, personalized therapeutic strategies.

## 21.2 Complex clinical predisposition with complex complications

In 1859 Samuel Wilks, at Guy's Hospital in London, differentiated UC from bacterial dysentery. By 1931 Sir Arthur Hurst described clinical and sigmoidoscopic features of the disease and showed that the rectum was always affected, with extension proximally to involve a variable length of the colon. The disease is most commonly seen in individuals between 20 and 40 years of age, but it can be present in the first few months of life or in adults in their eighties. The disease is characterized by a relapsing and remitting course in the majority of patients.

The formal recognition of CD as distinct from UC had to await the finding in 1932 by Crohn, Ginzburg, and Oppenheimer. The granulomatous (aggregation of macrophages together with T lymphocytes) nature of the disease was recognized, but only in approximately 65% of patients, and it soon became clear that the disease could affect any part of the gastrointestinal tract. Nevertheless, it is by no means clear that UC and CD are two homogeneous disease entities (Table 21.1). The discovery of new biomarkers could facilitate the noninvasive differentiation between the subtypes of CD -colonic, and ileocolonic. When CD is confined to the colon, it may be very difficult to distinguish it from UC in 5%–10%—these patients are currently labeled “colitis not yet classified.” Colonic CD and UC patients are both associated with a greater overall relative risk of colorectal cancer [1–3]. In mice affected with colitis the

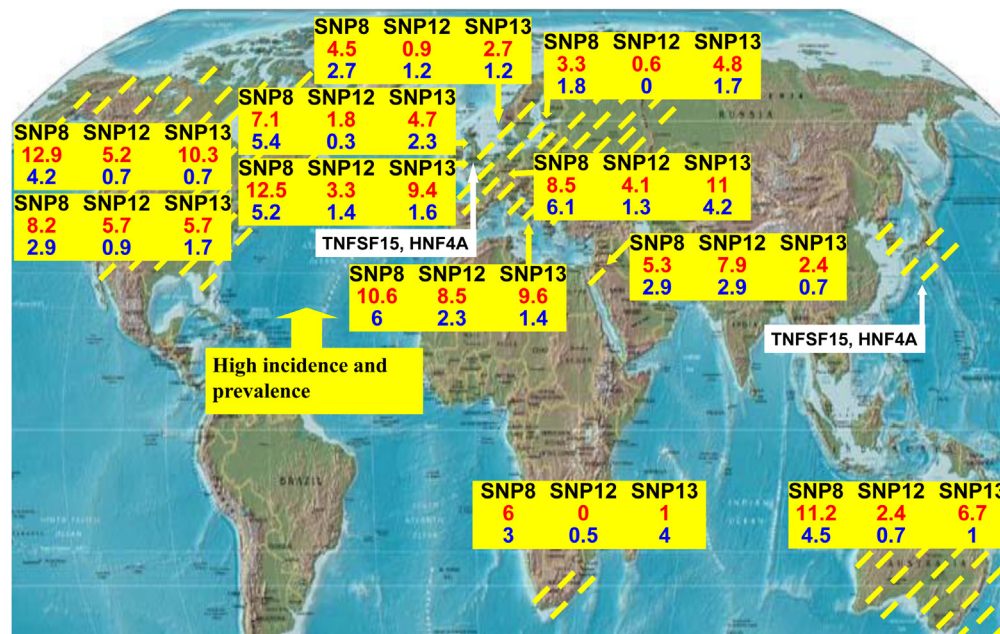
**TABLE 21.1** Comparison of Crohn's disease with ulcerative colitis.

	Crohn's disease	Ulcerative colitis
<b>Genetic epidemiology</b>		
Prevalence	27–106 per 100,000	80–150 per 100,000
Incidence	4–10 per 100,000	6–15 per 100,000
Sibling affected risk ratio	30–42	4
Concordance between identical twins (%)	37	10
Concordance between nonidentical twins (%)	10	3
Mode of inheritance	Complex	Complex
Loci predominantly linked with either Crohn's or ulcerative colitis	<i>NOD2</i> , <i>ATG16L1</i> , <i>IRGM</i>	<i>HNF4A</i> , <i>LAMB1</i> , <i>CDH1</i> , and <i>GNA12</i>
MHC	HLA DR7, DRB3*0301, and DQ7 (associated with colonic Crohn's)	HLA class II phenotypes DR2, DR9, and DRB1*0103
Predisposing environmental factor	Smoking	Smoking
Protective environmental factor		
<b>Clinical features</b>		
Bloody diarrhea	Less common	Common
Abdominal mass	Common	Rare
Perianal disease	Common	Less common
Malabsorption	Frequent (ileal disease)	Never
<b>Radiological/Endoscopic features</b>		
Rectal involvement	Frequently spared	Invariable
Distribution	Segmental discontinuous	Continuous
Mucosa	Cobblestones, fissure ulcers	Fine ulceration, “double contour”
Strictures (narrowing)	Common	Rare
Fistulas (opening)	Frequent	Rare
<b>Histological features</b>		
Distribution	Transmural	Mucosal
Cellular infiltrate	Lymphocytes, plasma cells, macrophages	Lymphocytes, polymorphs, plasma cells, eosinophils
Glands	Gland preservation	Mucus depletion, gland destruction, crypt abscess
<b>Special features</b>		
	Aphthoid ulcers, granulomas, histiocyte-lined fissures	None

HLA, Human leukocyte antigens; MHC, major histocompatibility complex.

From Dhavendra Kumar, David Weatherall, *Genomics and Clinical Medicine*, Oxford University Press, 2008.





**FIGURE 21.1** The regions known for high incidence and prevalence of inflammatory bowel diseases are highlighted on the global atlas. The allele frequencies (shown in red with controls in blue) indicate association of the heterozygote caspase recruitment domain family, member 15 (CARD15) mutations Arg702Trp (SNP8), Gly908Arg (SNP12), and the frameshift 3020 insC (SNP13) among populations of European descent and European admixture populations affected with Crohn's disease (see the "Epidemiology of CARD15" section; Hampe et al., 2002; Cavanaugh et al., 2003; Fidder et al., 2003; Palmieri et al., 2003; Zaahl et al., 2005). In contrast, a core *TNFSF15* haplotype is associated with CD in both Caucasians and Japanese populations (see the "TNFSF15" section). Furthermore, the *HNF4A* locus is also associated with UC in both Caucasian and Japanese populations. *CD*, Crohn's disease; *UC*, ulcerative colitis.

increased risk of cancer can be due to the toxicity of the altered microbial composition [4] as well as modulation by interleukin 22 (IL-22), a cytokine of the IL-10 superfamily [5].

### 21.2.1 Epidemiology

Over the last century, CD and UC have manifested an increasing incidence in various global populations, with a current estimated prevalence of 0.15% in the Northwest Europe and North America [6]. The environmental influences associated with modern urban developments have coincided with increasing incidence and prevalence of IBD in industrialized countries (see Fig. 21.1). The clinical manifestation of IBDs is likely to be due to a provocation of the mucosal immune system by commensal intestinal flora in susceptible individuals—in most individuals, these would be nonpathogenic [7]. Accumulating evidence from genetic epidemiology implies that the host genome predisposes to autoimmune and/or inflammatory diseases [8]. However, genetics alone is not adequate in explaining the disease, since animal models that are genetically altered for predisposition to colitis do not develop the phenotype when kept in a germ-free environment [9,10]. The impact of early life exposure to microbes is crucial for dampening the symptoms of IBD and consistent with the hygiene hypothesis [11,12].

### 21.2.2 Genome versus environome

The inheritance of IBD was observed prior to the application of genomics and molecular systems biology being envisioned. This included the manifestation of familial clustering as observed by Kirsner in 1963 and thus set the foundation for subsequent genetic investigations. Furthermore, within families, there was a striking concordance of clinical characteristics [13,14]. A more rigorous validation was provided from twin studies that compared disease concordance in monozygotic (MZ) and dizygotic (DZ) twins, because familial clustering does not entirely exclude environmental influences [15–18]. The twin studies provide an indication of the maximum risk of occurrence. The concordance rates among MZ twins with CD and UC range, respectively, from 37% to 10%, while the corresponding concordance rates among DZ twins are only 10% to 3%. Higher concordance rates are seen in CD



than in UC twin pairs. Since twins in general share similar environments during childhood, the higher disease concordance rates in MZ twins are a very strong argument for genetic susceptibility.

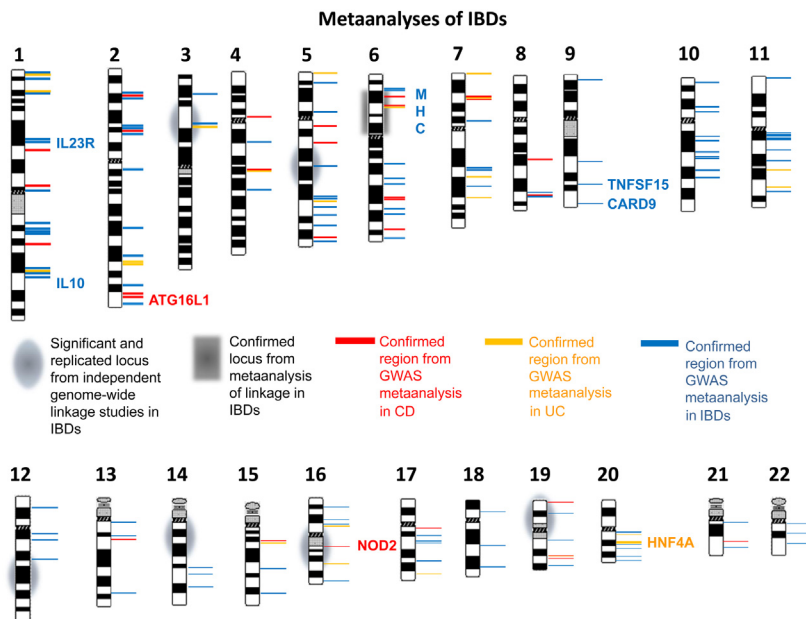
An alternative to a comparison of the concordance rates between MZ and DZ twins affected with IBD is the measurement of risk-to-relative ratio for a relative of type R, known as  $\lambda_R$ ; this method also yields an assessment of the genetic component [19]. The  $\lambda_R$  quantifies the familial component of a discrete trait that describes the relationship: first-degree siblings, closely related family members, and/or other relatives. Therefore if the risk for a relative or affected sib is 3%–3.5% for CD, and the population prevalence is 0.1%–0.2%, then the  $\lambda_S$  value is 15–35 [20]. The value between 6 and 9 for UC [21] implies that the genetic influence is not as powerful as it is for CD. This measurement could include discrepancies due to the introduction of a bias if the family members are exposed to the same environmental factors [22]. The  $\lambda_S$  value may fluctuate among different populations. Second-generation immigrants, unlike the first-generation, have a similar incidence of IBD compared to that of the nonimmigrants [23]. This implies that there is evidence for the “environome” having a major influence. The environome includes the diet, antibiotics, and immunosuppressives, as well as the microbiome (a representation of the genomes of all of the microorganisms present within the gastrointestinal tract) [11,24]. The environome can perturb the microbiota (all of the microorganisms, such as bacteria, archaea, fungi, and viruses) and lead to dysbiosis (a reduction in bacterial diversity) in humans, such as a decreased proportion of *Faecalibacterium prausnitzii*, and has been associated with IBD [25]. Subsequently, this could be followed by an increase in the complexity of bacteriophages [26]. In addition to bacteria, the microbiota can also include fungi, protozoa, and helminthes. *CARD9*, a pattern-recognition receptor (see *CARD9*), associated with IBD, *CARD9* deficiency was also altered with mycobiome (fungal) diversity [27]. It has also been hypothesized that an infectious agent could cause the common IBD clinical phenotype [28], as is the case for exceedingly rare primary immunodeficiencies that resemble unusual forms of IBD [29] (see parallels and paradoxes). Other factors that could be considered the environome are vitamin D [30], smoking [31], and dietary intake, such as omega-3-polyunsaturated fatty acids [32,33].

### 21.2.3 Genetics and genomics

Genetics and genomics have a key role for the molecular systems approach. The hypothesis-driven genetic studies have been successful in a limited number of the cases (see *NOD2* and *MHC*). The top-down approaches of using nonhypothesis genomic methodologies have provided the proof of principle for the identification of both polygenic variants associated with IBD and the inheritance of the monogenic causing disease mutations.

IBD phenotypes have been identified that can segregate in families [34–37]. Segregation analysis followed by homozygosity mapping with genetic markers, such as the microsatellites, and subsequently sequencing the candidate genes, such as the *IL10R* and *IL10*, has implicated mutations for the oncogenic fois [38,39]. These patients from consanguineous families manifested an autosomal-recessive severe enterocolitis within the first year of life. In addition to these exceedingly rare mutations, relatively rare polymorphisms within the *IL10R* and *IL10* genes have also been associated with very early onset of UC [40]. Originally, the *IL10* association was detected within a European UC case-control study [41], and subsequently it has been confirmed within a broader spectrum of IBD from a number of different geographical populations [8]. The *IL10* and other monogenic forms of the disease [34–37] could shed light on the common polygenic IBD (see parallels and paradoxes).

The IBD phenotype tends to aggregate in families rather than segregate within families, as is the case for the monogenic forms of the disease. Linkage analysis with evenly spaced microsatellite markers has also been applied in a collection of families, followed by denser linkage mapping and further fine mapping by association with single-nucleotide polymorphisms (SNPs) and other microsatellite markers. Such approaches led to the identification of *NOD2* at 16q12 (see *NOD2* section; [42]). Historically, the genome-wide linkage studies were undertaken to map genes that would otherwise not to be considered, as well as for enabling the coverage of the exhaustive list of candidate genes and loci that require systematic analysis. There have been up to a dozen of independent genome search results that have been reported from multiple affected IBD families of European ancestry, and an emerging pattern of replication now reaffirms that the following loci are significantly linked with IBD (Fig. 21.2): 16q, 12q, 6p, 14q, 5q, 19p, 1p, 16p, and 3p [44,45]. Furthermore, two metaanalyses by pooling the results from the different genome searches have shown that 6p is significantly linked with disease, and that 16q and 19p show suggestive linkage [43,46]. The reasons for these loci not being consistently replicated in all of the independent investigations could be due to one or more of the following: a false positive; a true locus that is population specific; a study with low power; a variable number of CD and UC patients; different criteria for



**FIGURE 21.2** The metaanalysis of GWAS has confirmed 110 (blue) IBD loci, and 30 CD-specific (red) and 23 UC-specific loci. Linkage to IBD at the MHC locus (6p) was also confirmed by a metaanalysis. CD, Crohn's disease; GWAS, genome-wide association studies; IBD, inflammatory bowel disease; MHC, major histocompatibility complex; UC, ulcerative colitis. Source: Adapted from van Heel DA, et al. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum Mol Genet* 2004;13(7):763–70.

diagnosis; different markers; and variation in genotyping quality. Similarly, pooling the data for metaanalysis can be hindered by these differences between the different investigations.

Unlike the SNPs within the *NOD2* gene that were associated with CD [8], the exceedingly rare disease-causing mutations, such as that within the *IL10* and *IL10R* genes that segregated with a very extreme phenotype, were not identical by descent between different families [38,39]—even if the defective gene was the same. Therefore family studies using linkage analysis, rather than population case-control studies, are more appropriate in assessing contribution to disease, because the very rare disease allele that is associated with the disease would be more likely to occur in certain families rather than that at the level of a population, and it would not be sufficient to lead to association at the level of the population. Thus although linkage analysis is applicable for an exceedingly rare disease-causing mutation, a population case-control study for association would not be appropriate for some of the very rare mutations; for example, those described within the *IL10* and *IL10R* genes in colitis [38,39].

Genome-wide association studies (GWAS) marked a new era for case-control investigations with SNP markers in CD, as a successor to genome-wide search with dinucleotide repeat sizing for linkage analysis in families. The prelude to the prolific nature of SNP genotyping was highlighted by the identification of the association between the tumor necrosis factor ligand superfamily member 15 (*TNFSF15*) gene at 9q32 and 94 CD patients compared to 752 healthy controls in a Japanese study [47]. For fine mapping and GWAS, the availability of a higher density of SNPs has offered a higher resolution and performs better at detecting association in a region of short-range linkage disequilibrium LD [48]. Furthermore, in populations of European descent, this locus has been confirmed in IBD [8]. Thus in contrast to *NOD2* mutations and the IBD5 haplotype, a core *TNFSF15* haplotype associated with IBD is found in both Caucasian and Japanese populations (see Fig. 21.1).

Direct genotyping variants with potential for function, such as nonsynonymous genome-wide SNPs, has successfully implicated an autophagy-related gene (see Autophagy section further on), the autophagy 16–like gene (*ATG16L1*) at chromosome 2q37 in a German population case-control investigation [49]. Microarray platforms that utilize genome-wide haplotype-tagged SNPs derived from the HapMap project have also efficiently led to the identification of novel genetic variants that are associated with IBD. This was exemplified by the implication of a subunit of the receptor for the proinflammatory cytokine interleukin-23 (*IL23R*) gene at chromosome 1p31 with noncoding variants that increased the risk of disease, and a rare coding variant that protected from disease in North American populations of European ancestry [50]. The other novel genome-wide significant associated regions from an extension to this investigation included paired-like homeobox 2B (*PHOX2B*) 4p13, neutrophil cytosolic factor 4 (*NCF4*) at 22q12, a predicted gene (*FAM92B*) at 16q24, an intergenic locus at 10q21, as well as additional evidence for *ATG16L1* [51]. A similar approach in a Belgian investigation discovered a novel association of CD with a gene desert region with no recognized genes other than CpG islands on chromosome 5p13 [52], and expression quantitative trait loci (eQTL) analysis revealed the disease-associated polymorphism that regulated the expression of the *PTGER* gene. A larger scale GWAS with high-density randomly genotyped SNPs

in the United Kingdom had independently unraveled another novel autophagy-inducing gene, the p47 immunity-related GTPase (*IRGM*) gene on chromosome 5q33 [53,54] (also see CNV, and Autophagy in the next section). This investigation also revealed a novel association with gene deserts on chromosome 1q. The other novel loci included 3p21, 5q33, 10q24, and 21q22. Also, the analysis of the results of GWAS among several diseases showed a novel association with the T-cell protein tyrosine phosphatase (*PTPN2*) gene locus at chromosome 18p in both CD and type 1 diabetes (T1D), which subsequently was replicated in both diseases [53,55]. There was a consistent pattern of replication between the different GWAS, but the intersection between genome-wide linkage and association was only visible at the 16q and 5q loci (see Fig. 21.2). This correlation also seemed to occur at 6p and 3p—these loci are gene-dense and show higher levels of LD that make it difficult to identify the causative variants. However, the other loci show no correlation between the previously undertaken genome-wide linkage scans and the recently undertaken GWAS. The meticulous genotyping of microsatellite markers can yield equivalent information in comparison to that of a higher density of SNP markers for linkage analysis, thus their use might not have been a major reason for the lack of correlation between genome-wide linkage and association studies. In retrospect, an overoptimistic estimation of weak replication between genome-wide linkage scans to reaffirm loci might be one of the contributing factors that led to a lack of correlation between the more robust GWAS. In addition, the regions of aggregation in linkage are broader, and their boundaries are not precisely demarcated.

The first generation of commercially available SNP microarray platforms had a similar overall coverage, but the coverage was incomplete in different regions of the genome [56]. Several different systematic metaanalyses and follow-up studies have combined the results from the different GWAS platforms and new cohorts and refined the locations of overlap between the IBD loci [8,57–62]. The large metaanalysis and fine mapping of IBDs has included a total of over 75,000 cases and controls derived, in addition to a follow-up with a customized high-density fine-mapping array known as the “ImmunoChip” [8,63]. The high-density markers for the ImmunoChip were selected from the associated regions of a number of different immune diseases. These data presented substantial evidence for the sharing of biological pathways with other diseases, such as inflammatory diseases, autoimmune diseases, mycobacterial infections, and primary immunodeficiencies. The shared associations with other immune-mediated diseases suggested that infectious diseases are likely to be the strongest selection pressures. The efficient genotyping of surrogate SNPs from the “HapMap3” reference set enabled the prediction of 1.23 million genotypes, and the assigned genotypes within IBD GWAS regions were subsequently validated by further genotyping with the ImmunoChip. The two major phenotypes of IBD—CD and UC—have more than 110 loci that are shared between them. More specifically, at least 30 are associated more CD than UC; and at least 23 more UC than CD.

Other markers, such as trinucleotide repeats, might be of direct relevance in IBD susceptibility, if the expansion of their repeat size is associated with disease [64], similar to the “anticipation” that has previously been described for Huntington’s disease [65]. There is, however, no conclusive evidence regarding the role of anticipation in IBD, since this observation could merely be an earlier diagnosis in the later-generation members of a family who are under clinical investigation, rather than genuinely earlier onset of disease than in the preceding generation [66]. Long-read sequencing could enable the reevaluation of the role of trinucleotide repeat markers [67].

Further insights of the contribution of copy number variants (CNVs) in IBD have been gained from the identification of a deletion within the *IRGM* gene after fine mapping a tagging SNP marker that was in perfect LD with a deletion CNV [68]. The deletion polymorphism was found to influence the expression of the *IRGM* gene. Subsequently, a reevaluation of the genome for CNVs suggested that, except for the CNV at the *IRGM* and human leukocyte antigens (*HLA*) genes, the other common CNVs are not likely to be associated with IBD [69].

The IBD-associated SNP markers explain less than 30% of the estimated inheritance, and further GWAS and metaanalysis of IBD are not likely to discover many more than the 200 distinct loci that have already been discovered [8,63,70,71]. Therefore the “missing inheritance” might be due to a number of different reasons. These include the possibility that an overestimation of the genetic component from the twin studies was due to previous methodological limitations [72]. Undetected monogenic mutations may also have a reduced penetrance for the complete syndrome [29]. The rare variants that have a modest influence [odds ratio (OR): 1.1–1.5] are not likely to capture the association by case-control studies [73,74]. There is also the possibility that a single-marker model that assumed complete LD could have missed some of the distinct genetic contributions, such as another distinct gene association, *CYLD*, in addition to the *NOD2* at 16q [75]. Another reason for the “missing inheritance” from the GWAS could be that gene–gene interactions are computationally challenging to assess and require very large sample sizes [76]. Pathway-based analysis of GWAS markers with a prior biological understanding of gene function is likely to have greater power to detect associations [77].

Although the high-resolution SNP markers can be relatively abstract for mapping, they have been pivotal to painting the genomic portrait of polygenic IBD [63]. However, as an extension to the deep resequencing of the GWAS-associated regions [78], the whole genome resequencing at base pair resolution could be an ultimate entity for single bp mapping [63,70] and can be combined with complementary methods of Bayesian fine-mapping methodologies [63]; and, although it may not be a complete blueprint for IBD, it is likely to be a starting point for unraveling IBD's complexity. The pioneering integration of sequencing, proteins, and metabolic pathways has begun to correlate the microbiota signatures of CD [79]. It has also been hypothesized that in addition to the core genes of the disease pathways, almost any of the other the genes outside the cluster of the biological central disease pathways can accumulate with a tiny influence toward the disease, known as the "omnigenic" model [80]. Genetics can also influence the gut microbiome [81–83]. Multiomics studies have so far demonstrated the interaction between of genomics, epigenomics, proteomics, metabolomics, and the microbiome in the regulation of the immune system [84,85].

#### 21.2.4 Epigenome

Historically, epigenetics referred to the interactions between the inherited factors and the environment during development [86]. Currently, epigenetics and epigenomics are more broadly used to describe the mechanism of gene regulation associated with phenotypes, such as IBD, that are not due to changes in the DNA sequence even though they can sometimes be inherited [87]. Epigenomics may also reveal part of the missing inheritance [88,89]. Moreover, epigenetics may mediate between the influences of an altered composition of the microbiome and the host genome [90].

The epigenomic mechanisms include DNA methylation (biochemical modification), histone modifications, noncoding RNAs, autoregulatory proteins, and chromatin interactions. Of the different epigenetic mechanisms that could modulate IBD, DNA methylation has been the most widely examined and could be considered an IBD risk factor [91,92]. GWAS have revealed association between CD and epigenetic regulatory enzymes such as DNA methyltransferase (DNMT) 3a [62], and subsequently DNMT3b [8]. Whole-genomic assessment of methylation by sequencing at nucleotide resolution can be undertaken by shotgun sequencing of bisulfite-converted DNA (MethylC-seq) [93,94]. Originally, epigenome-wide methylation association studies were undertaken with platform-based arrays, in tissues, such as peripheral blood samples [88]. This approach in peripheral blood has revealed 50 genes being differentially methylated in CD [95]. This epigenetic modifications have also included immunity-related pathways or cytokine-mediated process [88,96,97]. Methylomic and transcriptomic profiles have also been integrated from intestinal biopsies [98]. Chromatin accessibility combined with gene expression, localization of the inflammation, genotype, and the microbiome composition has suggested new subtypes of CD [99,100]. Chromatin modifications can modulate gene expression as a fibrosis of the intestine during an irreversible complication of IBD, the fibrosis of the intestine [97]. Chromatin immunoprecipitation can enable the localizing of chromatin modifications [101,102], followed by the application microarrays, or increasingly by sequencing (ChIP-seq) [103,104]. Epigenetics can provide insights for gene expression in different innate immune (see innate immunity) cell types [105]. The protein–DNA interactions, as investigated by ChIP-seq, can be combined with the quantitative measurements of the transcriptome [106].

#### 21.2.5 Transcriptome

The biological context of the GWAS in tissues has been derived from the entire set of RNA or transcriptome within relevant cells [107,108]. Unlike genomics, transcriptomics is dynamic and can fluctuate in response to different conditions. The heterogeneity can be further reduced by the identification of new of the cell subtypes [109]. This is the expression profile of RNAs in a specific cell or tissue. RNA sequencing by cDNA and next generation sequencing (NGS) is a sensitive method for genome-wide transcription factor binding-site profiling and transcription characterization that is called RNA-Seq [110]. The whole-transcriptome sequencing with RNA-Seq also yields quantitative results. This technique can measure the absolute quantity and capture transcriptome dynamics across distinct systems without extensive normalization of data sets. The dynamic manifestation of the genome can be yielded from the entire RNA profile. This can be viewed from both a holistic approach within the gastrointestinal systems and by reducing the heterogeneity by investigating the transcriptome at the level of the single-cell subtype [109]. The noncoding variants that are associated with disease were observed to have



regulated the eQTL [107]. The eQTL analysis has been informative for fine mapping the causative gene such in the major histocompatibility complex (MHC) (see Section 21.3), especially in regions of high LD.

Gene expression levels in both the pathogen and the host can simultaneously be analyzed by Dal RNA-seq of pathogen and host [111]. Dissection of the IBD transcriptome has originally been undertaken by using genome-wide cDNA microarrays. The transcriptional signature that coincides with mucosa of UC patients in remission was significantly different despite healing [112] (see therapeutics). Mucosal gene expression differentiates UC from colonic CD in pediatric patients during treatment [113]. A transcriptional risk score can also be derived from the risk-allele-associated gene expression and could be more informative than the genotypic risk assessment [114].

Other IBD-associated biomarkers, such as microRNAs (miRNAs) can be included for their expression signatures and can also be correlated between the different subtypes of CD and UC [115–117]. miRNAs are 21–23 nucleotides in length, which are evolutionarily conserved and are able to regulate the expression of protein-coding genes and are stable in tissues and biological fluids and therefore their potential as diagnostic or prognostic biomarkers [118].

### 21.2.6 Proteomics and metabolomics

The proteins are proximal to the phenotype, unlike the genome, and more than the transcriptome. Furthermore, there is not always a direct relationship between the levels of a transcript and its corresponding protein [119] due to posttranslational modifications [120], and subsequently an amplification of protein structure and function. Proteomics is the high-throughput assessment of the diversity of the proteins within the host or the microbiota and can be included for investigating the interactions between metabolism, cellular networks, and the environment. Proteomics is of central importance for the discovery of new biomarkers and measuring enzymatic activity [121].

The pioneering application of proteomics in IBD has included monitoring the protein modifications in intestinal cells due to different cytokines [122]. Another investigation has revealed the proteomic profile of the intestinal mucosa in UC [123]. Machine learning tools combined with proteomics could also enable enhanced differentiation between the different subtypes of CD and UC [124]. The metabolomic signature can be investigated in fecal extracts, plasma [125], serum [126], colonic tissues [127], and in urine [128].

The metabolic constituents of the cell, the metabolites, can highlight the individuality about the cellular state [129]. This approach has suggested that the metabolic state is altered more in CD than UC [130]. Noninvasive biomarkers from metabolite profiling, such as fatty acids, in feces, urine, and blood samples could be correlated with disease-related progression of IBD [131].

---

## 21.3 The identification of the NOD2 gene

The initial genome-wide linkage analysis search for IBD susceptibility made a great leap forward with the semiautomated genotyping of 270 fluorescently labeled dinucleotide-repeat microsatellite markers in 78 multiple sibs affected with CD [132]. This model-free approach identified a novel significant linkage at chromosome 16q. Subsequent to the completion of other genome searches, it was found that this locus was much more consistently replicated than were other regions that were reported to confer susceptibility to IBD [46,133,134]. Furthermore, the locus-specific sibling locus was estimated at 1.3 as predicted by the proportion of expected allele-sharing by affected sib pairs identical by descent (i.e., 25%) to the observed allele-sharing (19.2%) [132]. From this locus value, it was calculated [135] that the 16q might contribute less than 20% of the genetic influence in CD [136]. Linkage of the disease to 16q has been confirmed by two metaanalyses of published genome-wide screens [43,46]. This seminal discovery is regarded as the first novel, unequivocal identification of a susceptibility gene using nonparametric approaches [42]. Equally remarkable were the results from the positional candidate-gene approach that was undertaken simultaneously by an independent research group [137] and subsequently replicated [138] the rationale for selecting the *nucleotide-binding oligomerization domain containing 2* (*NOD2*) as a positional candidate-gene related to functional investigations that were emerging for the role of this gene in innate immunity [137]. Deep sequencing has also enabled the identification of other relatively rare variants within the *NOD2* genes that are associated with CD [78]. The successful genetic incrimination of *NOD2* gene in CD has opened new questions of how the mutations cause disease and their implications for the epidemiology of disease. Other *NOD*-like receptors could also be considered candidates for influencing CD [139].



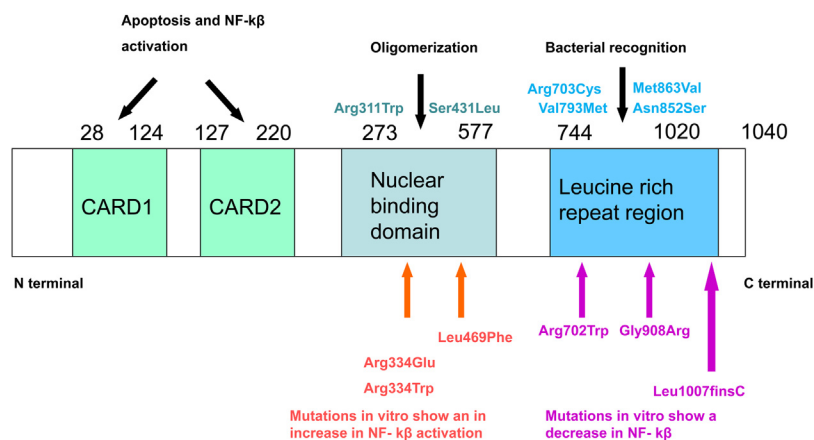
## 21.4 NOD2 and innate immunity

The *NOD2* gene belongs to a family of cytosolic pattern-recognition receptors [140]. Its role in innate immunity is based on its ability to recognize conserved structures within the gut flora [141]. There are now more than 20 different *NOD*-like human proteins, which show varying degrees of homology to plant cytosolic R proteins. An acronym puts forward to describe *NOD*-like human proteins is CATERPILLAR, because leucine-rich repeats are found in the carboxy termini [CATERPILLAR: CArd, Transcription Enhancer, R (Purine)-binding, pyrin, and many Leucine repeats] [142].

The *NOD2* gene consists of a central nucleotide-binding domain (NBD), a leucine-rich repeat, and two N-terminal CARD domains. These three common variants of *NOD2* (Fig. 21.3) are located within the leucine-rich repeat; this includes the Arg702Trp (tryptophan substituted for arginine at codon 702), the Gly908Arg (arginine substituted for glycine at codon 908), and Leu1007finsC (frameshift mutation truncates the ending 3% of protein). Subsequently, a number of relatively rare SNPs and many other private mutations in the *NOD2* gene have been found in patients with CD [78]. Other novel, rare mutations in the NBD were found to segregate with Blau syndrome, an autosomal granulomatous disease [143]. These mutations in the *NOD2* gene have been shown to provide a gain of function with increased nuclear factor- $\kappa$ B (NF- $\kappa$ B) activation in Blau syndrome [144]. However, not all granulomatous diseases are associated with *NOD2* mutations, as they were not found in sarcoidosis [145,146], although association has been found in patients with early onset of disease [143].

*NOD2* is an intracellular pattern-recognition protein for bacterial detection [147]. The recognition of the peptidoglycan component of bacterial cell walls depends on the detection of muramyl dipeptide (MDP), a minimal motif that is an almost-universal constituent of peptoglycans found in Gram-negative and Gram-positive bacteria [141,148]. Specifically, it is the leucine-rich repeat region of *NOD2* that has a role in binding to MDP, and consequently activating NF- $\kappa$ B, through a number of other intracellular molecules [140,147,148]. Paradoxically, in vitro experiments have shown that common variants of *NOD2* associated with CD decrease NF- $\kappa$ B activation [149,150]. It was hypothesized that this contradiction between in vitro and in vivo findings might be explained by *NOD2*-independent bacterial activation of NF- $\kappa$ B mediated by Toll-like receptors on macrophages [151]. Experimental results from in vivo experiments have been consistent with the hypothesis [152–155]. Further insights have also come from the transfection of epithelial cell lines with the CD-associated variants of *NOD2*. The transfected cells failed to kill intracellular *Salmonella typhimurium* compared to cells transfected with wild-type *NOD2* [156]. Another mechanism whereby *NOD2* mutations may put individuals at risk of disease has been proposed. The expression of *NOD2* has been shown in human Paneth cells [157]. Subsequently, CD patients with *NOD2* mutations have been shown to have a reduced release of the antimicrobial peptides,  $\alpha$ -defensins 4 and 5, from Paneth cells [158–163]. Since Paneth cells are predominantly found in the terminal ileum, deficiency of antimicrobial peptides in the intestinal lumen may allow bacterial-induced inflammation to occur.

*Nod2* knockout murine models have also been used to address the inconsistency of MDP activation between in vitro and in vivo [164]. It was found that the detection of bacterial MDP was abolished in the absence of *Nod2* in mice. Furthermore, the *Nod2*-deficient mice were more susceptible than wild-type mice to infection with bacteria via the oral route, but not with intravenous or intraperitoneal (IP) routes of infection. However, the *Nod2*-deficient mice showed no evidence of intestinal inflammation. These mice also showed a reduced release of cryptdins from intestinal Paneth cells—cryptdins being the murine equivalent of human defensins.



**FIGURE 21.3** The three *NOD2* domains include the following: caspase recruitment, nuclear binding, and leucine-rich domain. The three relatively common CD-associated variants display a loss of function, but the rare mutations that are associated with Blau syndrome reveal a gain in function from in vitro studies. Other relatively rare CD-associated variants are also shown. *CD*, Crohn's disease.

In contrast to the loss of function in cell lines transfected with *NOD2* variants or the *NOD2* knockout models, knock-in of the *NOD2* frameshift mutation has shown a much more efficient induction of cytokine interleukin-1 $\beta$  and NF- $\kappa$ B in response to MDP [165]. Perhaps the presence of an alternative pathway for positive regulation of NF- $\kappa$ B activation in response to MDP could offer a possible reconciliation between elevated NF- $\kappa$ B activation in the experimental models using *NOD2* variants.

*NOD2* deficiency can alter the microbiome and result in colitis that is similar to that treated with dextran sulfate in which there is a large entry of luminal bacteria [152,166–168]. However, *NOD2* deficiency does not lead to CD-like inflammation in a spontaneous colitis model in which there is a limited entry of the luminal bacteria [152].

#### 21.4.1 Epidemiology of *NOD2* in Crohn's disease

The functional insights have been further supported by the compelling weight of the epidemiological disease association with populations of European descent [8,137,138,169–171]. These investigations have indicated that 10%–30% of CD patients are heterozygotes for one of the three common mutations (see Figs. 21.1 and 21.3). Another 3%–15% of patients are either homozygotes or compound heterozygotes. This relative risk of 20–40 in homozygotes (with a 1/25 risk of developing disease) is possible to estimate within a population case-control investigation. However, it is unlikely to influence the high-linkage score since homozygotes tend to reduce the log of the odds (LOD) score [172]. Furthermore, families unaffected by disease, yet possessing the high-risk alleles, have also been reported [170,173,174]. Moreover, within different populations of European descent, there is a variable association of these alleles with disease, even though the control allele frequencies within these populations are similar [175]. These differences in attributable risk of the *NOD2* disease-susceptibility variants could in some cases correlate with the apparent North–South gradient [176]. There is a reduced association of these susceptibility alleles with CD in Scotland [177,178], Ireland [179], and Finland [180]. Moreover, in northern Europe, the higher incidence and prevalence of disease is likely to be due to other genetic and environmental influences [15]. There is a notably reduced allele frequency of the disease-susceptibility alleles within Afro-American communities [181]. There is also an absence, or an exceedingly rare occurrence, of *NOD2* disease-susceptibility alleles within a number of Asian populations, such as the Koreans [182,183], Japanese [184,185], and Chinese [186,187]. In other communities, such as the Ashkenazi Jews, the much higher incidence and prevalence of CD could be irrespective of geographic and environmental influences in comparison to other groups [188]. In general, however, these differences of disease incidence and prevalence are usually thought to be due in part to migration and changes in the environment [189]. Although the much higher Gly908Arg allele frequencies within the Ashkenazi appear to correlate with disease [149], this particular genetic influence does not alone explain the much higher prevalence of CD within this community.

#### 21.4.2 *NOD2* mutations and phenotype

The association between the three common variants of *NOD2* has been consistently replicated with CD [8], but only a weak protective association with UC was noted [8]. Within CD [170,190], it was reported that the association was particularly strong for ileal CD. This has now been widely replicated and confirmed by metaanalysis within populations of European descent [175]. How *NOD2* variants influence the anatomical site of disease is unclear, but the possibility that Paneth cells are involved, as discussed in the preceding section, is intriguing. Paneth cells are predominantly located in the distal ileum, and it is possible that the lack of defensin production in individuals with *NOD2* mutations may impair ileal antimicrobial defense mechanisms. *NOD2* variants have also been reported to predispose to early onset of disease and to stricturing disease. However, these associations are not completely established, although they received some support from the metaanalysis [175].

#### 21.4.3 The Ancestor's tale of mutations that predispose to inflammatory bowel disease

The *NOD2* mutations that were associated with CD were likely to postdate the “out-of-Africa migration,” since this has so far not been detected in populations outside of those of European ancestry [184], and there is reduced allele frequency within the African-American communities [181]. The proline-to-serine amino acid substitution at position 268 of the *NOD2* gene has also not been detected in various global populations other than in those of European descent [184]. The substitution at position 268 has, however, been found in LD with all of the three common variants that predispose to CD in European populations [191]. Although the evolutionary significance

of this association is not understood [192], the absence of the variant allele at position 268 in various global communities other than in those of European descent suggests that it is not likely to precede the out-of-Africa migration. In addition, the LD between the variant allele at position 268 and the three other known disease-susceptibility alleles could have occurred by chance alone, since there is limited haplotype diversity within such a narrow region. Other, rarer mutations have also been identified within the *NOD2* gene [78]. Each of the different mutations may have a tale of its own about the protection from infectious diseases. The three most commonly CD-associated *NOD2* mutations were estimated to have arisen about 40,000 years ago [183,193], while the very rare *NOD2* mutations would have arisen more recently [78]. The age of the mutation and geographical distribution of allele frequencies yield clues to the survival advantage of this mutation. Mycobacteria are likely to have been a strong selective pressure for the IBD-associated SNPs [8].

Since the *CARD9* gene, a pattern-recognition receptor, associated with fungal infection in exceedingly rare immunodeficiency states [194,195], the relatively common fungal infectious diseases might have also been a selective pressure for the polymorphisms within the *CARD9* gene that were associated with IBD. Dectin-1 is also in the signaling pathway upstream from the adaptor *CARD9* [196], and the *Dectin-1* gene was associated with colitis patients' requiring surgery [27]. Furthermore, although previous research in mice models has indicated that an imbalance of commensal flora within gastrointestinal tract influences colitis, fungi may also contribute to colitis [27]. The complex haplotype structures of the MHC are also likely to have been under strong selection pressures from infectious diseases [8,197].

## 21.5 Major histocompatibility complex (6p21)

The MHC, commonly referred as HLA, mapped to chromosome 6 has been investigated over a period of 30 years. It has also been elaborated at a higher resolution. The molecular systems investigation of the MHC class I and class II has enabled further insights of the control and specificity of the antigen presentation [198]. The genetic region of the MHC is known to confer crucial immunological function and is linked with susceptibility to IBD [199–201] and many other autoimmune diseases. However, a good number of earlier investigations had yielded inconsistent results [202]. Unlike most other suspected autoimmune diseases, IBD has not shown the same level of consistent linkage to the MHC from independent genome-wide searches. In fact, it was only when a denser set of markers was applied in the MHC locus that it became possible to detect linkage [203]. The most consistent data within the *HLA* class II region have concerned *HLA-DR2* and *HLA-DRB1* 0103. The *DRB1* 1502 allele of *DR2* has been found predominantly in Japanese populations of UC patients [OR 3.74, confidence interval (CI) 2.20–6.38], but it has also been reported from North America and the United Kingdom [202,204–207]. The higher frequency of this allele in the Japanese general population, compared with its low frequency in Caucasians, probably explains the more consistent data from Japan. The association between UC and *HLA-DRB1* 0103 has been convincingly replicated in several large studies, which have also shown that patients possessing this allele are at risk of having severe colitis with a risk of colectomy [206,208,209]. It has been shown that this allele also influences the natural history of CD of the colon; that is, it is associated with severe disease and risk of colectomy early in the course of disease [210,211]. Whether this association with *DRB1* 0103 is causative or whether this allele is in tight LD with a more relevant allele is uncertain. However, *DRB1* 0103 appears to be present on only two small haplotypes, and thus a genuine association seems likely. The metaanalysis in 1999 also showed that there was a negative association between *HLA-DR4* and UC (OR 0.54, CI 0.43–0.68). Subsequently, this protective effect was shown to be due to the *DRB1* 0401 molecular subtype, but only when it is present on the *DRB1* 0401-DQB1 0301 haplotype [206,212]. For CD, metaanalysis has confirmed significant positive associations with *DR7* (OR 1.42, CI 1.16–1.74), *DRB3* 0301 (OR 2.18, CI 1.25–3.80), and *DQ4* (OR 1.88, CI 1.16–3.05), but, in contrast to UC, a negative association with *DR2*. The associations with *DR7* and *DR2* have been confirmed in a subsequent study [213]. Adjacent to the boundary between the *HLA* class I and class III regions, there are members of the nonclassical MHC class-I-related chain (MIC) gene family. *MICA* and *MICB* are polymorphic and are the only two members of the family to encode functional transcripts. They are predominantly expressed on the basal–lateral surface of epithelial cells and interact with *NKG2D* and a variety of natural killer cells, T cells, and macrophages, but particularly the *CD8 $\alpha$*  T cells and the  $\gamma$ T cells found in the intraepithelial compartment. So far, no consistent associations between UC or CD and polymorphisms in the *MICA* or *MICB* genes have been reported, and the most comprehensive study was negative [214]. Between the MHC class I and class II region is the highly gene-dense class III region, spanning approximately 900 kb. The genes for *TNFA*, *lymphotoxin  $\alpha$* , and heat-shock proteins are found within class III, and polymorphisms within these genes have been

associated with IBD. For *TNFA* a number of promoter polymorphisms have been studied (at positions 1031, 308, and 857). In the United Kingdom, *TNF*-857C was associated with CD patients not possessing *NOD2* mutations [215], while in Australia, it was present only in patients with *NOD2* variants [216]. Published data on the other promoter polymorphisms have been equally inconsistent [200,201]. The common promoter haplotype (*TNF* 1031T, 863C, 857C, 308G, 380G, and 238G) has been shown to be associated with distal UC, which is stable over time and, at least for a minimum of a 10-year follow-up, does not extend [206]. The mechanisms are unknown, but patients homozygous for this haplotype might have impaired TNF production well, which could influence disease activity. Polymorphisms in the *lymphotoxin a* gene and within a number of heat-shock protein genes are not consistently associated with either UC or CD (reviewed in [199]). Genes within the HLA region on Chr 6p may also determine whether patients with either UC or CD are at risk of developing extraintestinal manifestations. Thus the arthropathies, uveitis, erythema nodosum, recurrent mouth ulcers, and primary sclerosing cholangitis have been reported in association with class I and class II alleles [201]. For example, the reactive large-joint arthropathy seen in some patients with active colitis is strongly associated with HLA-B 27, HLA-B 35, and HLA-DRB1 0103, whereas the symmetrical small-joint arthropathy is associated with HLA-B 44 [217]. Uveitis has been associated with HLA-B27 and DRB1 0103, and erythema nodosum with *TNF*-1031C. The numbers are small, but it seems likely that the phenotypical heterogeneity seen in the clinic may partly be explained by genetic polymorphisms within the HLA region.

Association with the HLA locus has now been confirmed by the customized ImmunoChip and the metaanalysis [8]. In addition to the HLA genes at the Chr 6p locus, a protective association with a synonymous SNP Q350Q within the *BTNL2* gene has been consistently replicated with UC [218]. It is plausible that a synonymous SNP can influence disease [219]. A missense K196E *BTNL2* variant showed novel protective association in UC cases requiring colectomy, but the same allele was weakly predisposing for the ileal subphenotype in CD [218]. The subphenotypical studies require further replication.

## 21.6 The causative genome variants and functional implications

Of the over 200 IBD-associated variants, majority of the disease-associated SNPs were within the noncoding regions, and these including at least 63 that were correlated with gene expression [8,63]. In fact, even when deep sequencing was undertaken within the regions of the GWAS-associated genes, only a limited number of coding variants were identified, such as a *CARD9* splice variant [78]. However, even this relatively rare protective splice variant within the *CARD9* gene was distinct from the common *CARD9* predisposing variant that was previously reported within a CD GWAS metaanalysis [62]. The other notable coding variants that were identified from resequencing the IBD-associated GWAS regions include *IL18RAP*, *CUL2*, *C1orf106*, *PTPN22*, and *MUC19* [78].

The systematic interrogation of noncoding IBD-associated regions reveals the tendency to cluster within the regulatory DNA marked by deoxyribonuclease I (DNase I) hypersensitive sites (DHSs) [220]. About 88% of such DHSs have a role in fetal development. DHSs are likely to correlate with IBD associations that obstruct transcription factor-recognition sequences, frequently alter allelic chromatin states, and lead to the deregulation of networks in IBD. Transposase-accessible chromatin sequencing (ATAC-seq) [221] is increasingly being applied rather than having a more streamlined workflow and can be used significantly reduced cell numbers. not likely to define the causal variant when there is complete LD between a number of the plausible causative coding alleles, such as that between the associated alleles spanning one of the haplotypes between the *ORMDL3* and *GDSBM* genes at 17q21 [58]. The expression studies from Epstein-Barr virus-transformed lymphoblastoid cells from patients with CD and asthma have, however, presented complementary evidence, by showing that the coding SNPs with the strongest association were consistently associated with levels of *ORMDL3* transcripts rather than adjacent genes [58,222]. Further functional insights may also be possible from protein studies for the *ORMDL3* gene, which is a member of a conserved family of endoplasmic reticulum membrane proteins.

The macrophage-stimulating protein (MST1) modulates innate immunity and adaptive responses, and the coding variant R689C of the *MST1* gene was one of the likely causative variants within a region of high LD at 3p that was associated with IBD [223]. In addition, at 3p there has been some evidence for another distinct association within the receptor of biological *MST1* gene [59,224]. Since this region of association intersects with a region of linkage at 3p, it is likely that there are also several other rare SNPs that could influence the IBD [225].



## 21.7 Autophagy

Functional links have been established between innate immunity and autophagy [226,227]—autophagy is more literally known as “self-eating.” The microbial loads within infected cells are degraded within the lysosomes by the process of autophagy. This process recycles proteins and organelles for optimal cellular balance between synthesis and reducing the proteins into constituent components. Autophagy requires the formation of double-membrane cytosolic vesicles called *autophagosomes* that sequester cytoplasmic contents and transport them to the lysosome for subsequent degradation. The dynamic process of membrane expansion enables the disposal of microbes of any size.

The autophagy genes associated with CD, such as *ATG16L1*, *ATG5*, *IRGM*, and leucine-rich repeat kinase 2 (*LRRK2*) are expressed in the small intestinal Paneth cells. *NOD2* induces autophagy by directing *ATG16L1* to the cell membrane within the vicinity of bacterial entry [227]. The other autophagy genes, such as *RIPK2*, *ATG5*, and *ATG7*, are also involved. In addition to autophagy degrading the damaged organelles and proteins, it also connects with the process of antigen presentation. The bacterial handling and MHC class II antigen presentation occurs in human dendritic cells (DCs). The DCs of CD patients expressing CD risk variant *NOD2* or *ATG16L1* have impaired autophagy induction and reduced bacterial killing and defective antigen presentation [227]. This correlates with the major allele in cases having an impaired ability to capture bacteria.

*NOD2* also has a role in the autophagic response to bacteria engulfed by phagosomes [226]. CD-associated with the frameshift *NOD2* mutation as well as the *ATG16L1* variant results in defective bacterial breakdown in lysosomes [227]. Furthermore, *NOD2* and DCs can modulate Th17 responses [228]. The IL-23 signaling pathway stimulates antimicrobial peptide cryptidins in the intestinal cells, which would otherwise result in the impairment of bacterial destruction. Knockout mice for *ATG16L1* that are infected with the intestinal pathogen norovirus also develop defects in Paneth cells [229].

Colonic epithelial *IRGM* gene overexpression was observed in CD patients with a homozygous protective *IRGM* synonymous SNP allele, but not in patients with the risk allele [230]. In addition, the microRNA-196 (miR-196) was found to have an impaired binding to the allele that predisposes to CD. Therefore either an overexpression [230] or an underexpression of the *IRGM* gene [68] could reflect an impaired autophagy.

## 21.8 Adaptive immune system

The stimulation of adaptive immunity by the innate immune system results in a specific response with a greater effectiveness against infectious diseases than that of the innate immune system. Adaptive and innate immunity are, however, connected to each other in the most intimate way. In fact, adaptive immunity depends on innate immunity. *NOD2*, *IRGM*, *ATG16L1*, and *IL-23R* are examples of genes that regulate innate and adaptive immune responses. More specifically, for example, the activation of *NOD2* signaling serves to block the expression of IL-12, and indirectly IL-23, by the upregulation of a microRNA, miR-29 [231]. However, when there was a *NOD2* frameshift mutation, the miR-29 was not upregulated, which may explain how a mutated *NOD2* gene can lead to inflammation with elevated IL-23.

The *IL-23R* gene has one of the strongest associations with CD [8]. Furthermore, other genes within the IL23/Th17 signaling pathway, such as *IL12B*, *JAK2*, *TYK2*, and *STAT3*, were also associated with the disease. The results from the mouse model, which had also been studied extensively prior to the identification of genetic association, explained the biology underlying the genetic associations.

There are several other interconnected genes associated with CD, such as the *IL10*, *IL12B*, and *STAT3* [232]. This pathway contributes to the clonal selection of lymphocytes and sets the foundation for adaptive immunity with the expression of highly specific receptors on B and T cells.

## 21.9 Mucosal barrier function

For the UC-associated 20q locus, within populations of both Europeans [8,233] and Japanese [234], an impaired barrier function due to a defective *HNF4A* gene product is plausible. Mice models have also suggested that *Hnf4a* deficiency could enhance epithelial permeability, and an unbalanced *Hnf4a* gene expression may correlate with UC [235]. The other mucosal barrier genes that were specifically associated with UC include *ECM1* [57,236];



*LAMB1*, *CDH1*, and *GNA12* [8]. Glycoprotein mucin-2 (MUC2) can also function as a barrier between the gut microbiome and intestinal epithelium [237]. The defective mucous layer barrier can lead to severe colitis in mouse models [238]. This can modulate the gut microbiota and in turn contribute toward colitis [239]. Fucosyltransferase 2 (FUT2) deficiency at the mucosal barrier has also been associated with dysbiosis and IBD [240,241].

### 21.10 The parallels and paradoxes with other diseases

More than two-thirds of the distinct IBD associations transcend the boundaries between many other complex diseases that have a very different clinical disease manifestation [8,242,243]. However, there are several loci in common between UC and CD. In certain cases, there are emerging patterns of symmetry between different diseases and the shared loci. For example, the R620W *PTPN22* variant predisposes to T1D, systemic lupus erythematosus (SLE), and rheumatoid arthritis (RA), but it is protective for CD [8,54] and also weakly protective in UC [8].

Intriguingly, the *HNF4A* gene locus is associated with both maturity diabetes of the young (MODY) [244], the metabolic syndrome of type 2 diabetes (T2D; [245]), and UC. Both MODY and T2D have some clinical similarities. However, the *HNF4A* locus is not associated with T1D. The *HNF4A* gene is a plausible pleotropic gene. The other T2D and metabolic syndrome genes that may have parallels with IBD include *CDKAL1*, *CAPN10*, *GCKR*, *THADA*, *CPEB4*, *BTLN2*, *FADS1*, and *SMAD3* [242].

IBD is not likely to be a quintessential autoimmune disease such as SLE, ankylosing spondylitis (AS), or RA, yet there is a very high degree of shared features with the classical autoimmune diseases, such as the IL23R signaling pathway in and AS, SLE, and RA. The different autoimmune diseases may also cluster in affected patients [246]. Intriguingly, in Manitoba, Canada, there is a very high incidence of IBD in the population of European descent, but low in the First Nations. The difference in the frequencies for the alleles that are involved in bacteria processing may be one of the contributing factors for the protection from IBD in the First Nations [247].

The sharing of loci between IBD and infectious diseases are revelatory from the perspective of predicting the selection advantage of IBD variants. For example, a Chinese GWAS of leprosy (*Mycobacterium leprae*) indicated that the MHC genes were most significantly associated with leprosy. The *NOD2* variants in infectious diseases in Asian populations were distinct from the variants associated with CD in patients of European ancestry.

There are also shared loci between primary immunodeficiencies and IBD, such as the *CARD9* gene [194], which is paradoxical, given that an inflammatory disease is an overresponse rather than an underresponse, as is the case for an immunodeficiency. There has also been a report of a monogenic form of immunodeficiency and IBD being associated with a splice-site mutation in the *NEMO* gene on chromosome X [248,249]. Mutations in the *NEMO* gene result in the variable phenotype of incontinentia pigmenti (IP), an X-linked dominant disorder, predominantly seen in females, as it is lethal in males. Turner syndrome, a recognizable multiple-anomaly syndrome due to a number of constitutional X chromosome abnormalities, was also associated with IBD. More than 20 reported cases have shown this association [250]. The reported frequency of IBD has been found to be much higher in patients affected with Turners syndrome than in the general population. X-linked chronic granulomatous disease is also another monogenic immunodeficiency in which patients frequently manifest a granulomatous phenotype and Crohn's-like disease [251]. The precise mechanism by which primary immunodeficiencies and autoimmunity coincide remains to be unraveled, and the molecular systems approach are yet to be applied for the puzzle.

For polygenic granulomatous diseases, a predisposing association may be shared between complex *BTNL2*/*HLA* haplotypes, sarcoidosis [252], and primary biliary cirrhosis [253], and there may be parallels with ileal CD [218]. In UC, which is a nongranulomatous disease, these alleles were protective [218].

### 21.11 Clinical implications and translation

The potential to generate novel hypothesis from undertaking exploratory computational analysis with the massive multiomics data could enable the discovery of new biomarkers. The new biomarkers measured in accessible tissues are required for early diagnosis, monitoring disease progression, patient stratification, predicting recurrence, and therapeutic strategies. This includes the possibility of predicting disease severity in newly diagnosed patients with the transcriptional signature as a biomarker [242]. Furthermore, innovative approaches in proteomics may enhance the phenotyping of patients according to their pattern of protein expression [254].

The loci associated with CD prognosis (the course of the disease with time) are largely distinct from the GWAS CD [255]. GWAS have also revealed adverse drug response during therapeutic intervention [256,257]. However, a significant proportion of IBD patients do not respond to TNF neutralization at clinical presentation even though they are identical at clinical presentation to those that achieve remission [258,259]. The gut microbiome could also affect the efficacy of these immunotherapy [260,261]. Under different multiple “omics” conditions, different genes could be active [262].

Although these modestly influencing loci are currently not adequate to predict disease risk in unaffected individuals, even when combined with the transcriptome assessment [114], the identification of genes involved in autophagy and in innate immunity as well as adaptive immunity may provide greater insight into disease pathogenesis. The therapeutic intervention to activate autophagy in either Paneth cells or macrophages may decrease inflammation [263]. The detection of differences in circulating microRNAs (miRNAs) in the serum of IBD patients may also facilitate an earlier diagnosis [264], as well as a prognostic utility by integrating with messenger RNA [265,266]. The miRNAs may be able to exert an influence in about one-thirds of the genes in the human genome.

The discovery of the exceedingly rare *IL-10* and *IL-10* receptor polymorphisms that cause a severe phenotype has been followed up with novel treatments that include hematopoietic stem cell transplantation (HSCT). Allogeneic HSCT that restores *IL-10* signaling in hematopoietic cells seems to be a very promising new approach [232].

The early detection of tumor cells and the monitoring of circulating tumor cells in IBD patients with a predisposition to colorectal cancer are likely to develop with the emerging single-cell genomics technology [267].

## 21.12 Conclusion

Progress is emerging, though in its infancy, in the understanding of the molecular systems biology of IBDs, and this could add multiple dimensions to the understanding the multifactorial disease mechanisms of CD and UC. For all of the different OMICS approaches, the assembling of large cohorts of patients that have been meticulously documented in terms of multiple clinical phenotypes will be essential for the application of the multiomic technologies. The genetic knowledge so far has highlighted the important role of the host's innate immunity and its interaction with commensal bacteria, including autophagy, and the intimate integration between adaptive immunity, epithelial barrier function, and *IL-10* signaling. Further functional insights are likely to be revealed from the different OMICS technologies, and new hypothesis could be generated from the different molecular systems approaches; here is a sense of optimism that the molecular systems revolution will be translated into personalized therapy.

## References

- [1] Ekblom A, et al. Ulcerative colitis and colorectal cancer. A population-based study. *N Engl J Med* 1990;323(18):1228–33.
- [2] Ekblom A, et al. Increased risk of large-bowel cancer in Crohn's disease with colonic involvement. *Lancet* 1990;336(8711):357–9.
- [3] Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell* 2010;140(6):883–99.
- [4] Arthur JC, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 2012;338(6103):120–3.
- [5] Huber S, et al. IL-22BP is regulated by the inflammasome and modulates tumorigenesis in the intestine. *Nature* 2012;491(7423):259–63.
- [6] Binder V. Epidemiology of IBD during the twentieth century: an integrated view. *Best Pract Res Clin Gastroenterol* 2004;18(3):463–79.
- [7] Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007;448(7152):427–34.
- [8] Jostins L, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491(7422):119–24.
- [9] Strober W. Animal models of inflammatory bowel disease—an overview. *Dig Dis Sci* 1985;30(12 Suppl):3S–10S.
- [10] Blumberg RS, Saubermann LJ, Strober W. Animal models of mucosal inflammation and their relation to human inflammatory bowel disease. *Curr Opin Immunol* 1999;11(6):648–56.
- [11] Morgan XC, Segata N, Huttenhower C. Biodiversity and functional genomics in the human microbiome. *Trends Genet* 2013;29(1):51–8.
- [12] Olszak T, et al. Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* 2012;336(6080):489–93.
- [13] Satsangi J, et al. Clinical patterns of familial inflammatory bowel disease. *Gut* 1996;38(5):738–41.
- [14] Bayless TM, et al. Crohn's disease: concordance for site and clinical type in affected family members—potential hereditary influences. *Gastroenterology* 1996;111(3):573–9.
- [15] Halfvarson J, et al. CARD15/NOD2 polymorphisms do not explain concordance of Crohn's disease in Swedish monozygotic twins. *Dig Liver Dis* 2005;37(10):768–72.
- [16] Orholm M, et al. Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. *Scand J Gastroenterol* 2000;35(10):1075–81.
- [17] Thompson NP, et al. Genetics versus environment in inflammatory bowel disease: results of a British twin study. *Br Med J* 1996;312(7023):95–6.

- [18] Tysk C, et al. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* 1988;29(7):990–6.
- [19] Risch N. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet* 1990;46(2):229–41.
- [20] Satsangi J, et al. Genetics of inflammatory bowel disease. *Gut* 1994;35(5):696–700.
- [21] Orholm M, et al. Familial occurrence of inflammatory bowel disease. *N Engl J Med* 1991;324(2):84–8.
- [22] Guo SW. Inflation of sibling recurrence-risk ratio, due to ascertainment bias and/or overreporting. *Am J Hum Genet* 1998;63(1):252–8.
- [23] Pinsk V, et al. Inflammatory bowel disease in the South Asian pediatric population of British Columbia. *Am J Gastroenterol* 2007;102(5):1077–83.
- [24] Betteridge JD, et al. Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population. *Inflamm Bowel Dis* 2013;19(7):1421–7.
- [25] Cao Y, Shen J, Ran ZH. Association between *Faecalibacterium prausnitzii* reduction and inflammatory bowel disease: a meta-analysis and systematic review of the literature. *Gastroenterol Res Pract* 2014;2014:872725.
- [26] Norman JM, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 2015;160(3):447–60.
- [27] Iliiev ID, et al. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. *Science* 2012;336(6086):1314–17.
- [28] Segal AW. Making sense of the cause of Crohn's – a new look at an old disease. *F1000Res* 2016;5:2510.
- [29] Casanova JL, Abel L. Revisiting Crohn's disease as a primary immunodeficiency of macrophages. *J Exp Med* 2009;206(9):1839–43.
- [30] Cantorna MT, et al. Vitamin D, immune regulation, the microbiota, and inflammatory bowel disease. *Exp Biol Med (Maywood)* 2014;239(11):1524–30.
- [31] Beaugerie L, et al. Impact of cessation of smoking on the course of ulcerative colitis. *Am J Gastroenterol* 2001;96(7):2113–16.
- [32] Cabre E, Domenech E. Impact of environmental and dietary factors on the course of inflammatory bowel disease. *World J Gastroenterol* 2012;18(29):3814–22.
- [33] Cabre E, Manosa M, Gassull MA. Omega-3 fatty acids and inflammatory bowel diseases – a systematic review. *Br J Nutr* 2012;107(Suppl 2):S240–52.
- [34] Uhlig HH. Monogenic diseases associated with intestinal inflammation: implications for the understanding of inflammatory bowel disease. *Gut* 2013;62(12):1795–805.
- [35] Uhlig HH, Muise AM. Clinical genomics in inflammatory bowel disease. *Trends Genet* 2017;33:629–41.
- [36] Uhlig HH, et al. The diagnostic approach to monogenic very early onset inflammatory bowel disease. *Gastroenterology* 2014;147(5):990–1007 e3.
- [37] Worthey EA, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med* 2011;13(3):255–62.
- [38] Glocker EO, et al. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med* 2009;361(21):2033–45.
- [39] Glocker EO, et al. Infant colitis—it's in the genes. *Lancet* 2010;376(9748):1272.
- [40] Moran CJ, et al. IL-10R polymorphisms are associated with very-early-onset ulcerative colitis. *Inflamm Bowel Dis* 2013;19(1):115–23.
- [41] Franke A, et al. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet* 2008;40(11):1319–23.
- [42] Hugot JP, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;411(6837):599–603.
- [43] van Heel DA, et al. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. *Hum Mol Genet* 2004;13(7):763–70.
- [44] Ahmad T, et al. Clinical relevance of advances in genetics and pharmacogenetics of IBD. *Gastroenterology* 2004;126(6):1533–49.
- [45] Gaya DR, et al. New genes in inflammatory bowel disease: lessons for complex diseases? *Lancet* 2006;367(9518):1271–84.
- [46] Williams CN, et al. Using a genome-wide scan and meta-analysis to identify a novel IBD locus and confirm previously identified IBD loci. *Inflamm Bowel Dis* 2002;8(6):375–81.
- [47] Yamazaki K, et al. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum Mol Genet* 2005;14(22):3499–506.
- [48] Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 1999;22(2):139–44.
- [49] Hampe J, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007;39(2):207–11.
- [50] Duerr RH, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006;314(5804):1461–3.
- [51] Rioux JD, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;39(5):596–604.
- [52] Libioulle C, et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet* 2007;3(4):e58.
- [53] Parkes M, et al. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;39(7):830–2.
- [54] WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661–78.
- [55] Todd JA, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007;39(7):857–64.
- [56] Barrett JC, Cardon LR. Evaluating coverage of genome-wide association studies. *Nat Genet* 2006;38(6):659–62.
- [57] Anderson CA, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43(3):246–52.
- [58] Barrett JC, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40(8):955–62.

- [59] Beckly JB, et al. Two-stage candidate gene study of chromosome 3p demonstrates an association between nonsynonymous variants in the MST1R gene and Crohn's disease. *Inflamm Bowel Dis* 2008;14(4):500–7.
- [60] Cummings JR, et al. Contribution of the novel inflammatory bowel disease gene IL23R to disease susceptibility and phenotype. *Inflamm Bowel Dis* 2007;13(9):1063–8.
- [61] Cummings JR, et al. Confirmation of the role of ATG16L1 as a Crohn's disease susceptibility gene. *Inflamm Bowel Dis* 2007;13(8):941–6.
- [62] Franke A, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42(12):1118–25.
- [63] Huang H, et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature* 2017;547(7662):173–8.
- [64] Polito II JM, et al. Preliminary evidence for genetic anticipation in Crohn's disease. *Lancet* 1996;347(9004):798–800.
- [65] Vegter-van der Vlis M, Volkers WS, Went LN. Ages of death of children with Huntington's chorea and of their affected parents. *Ann Hum Genet* 1976;39(3):329–34.
- [66] Faybush EM, et al. Generational differences in the age at diagnosis with IBD: genetic anticipation, bias, or temporal effects. *Am J Gastroenterol* 2002;97(3):636–40.
- [67] Liu Q, et al. Interrogating the “unsequenceable” genomic trinucleotide repeat disorders by long-read sequencing. *Genome Med* 2017;9(1):65.
- [68] McCarroll SA, et al. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet* 2008;40:1107–12.
- [69] Wellcome Trust Case Control, Consortium, et al. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* 2010;464(7289):713–20.
- [70] Luo Y, et al. Exploring the genetic architecture of inflammatory bowel disease by whole-genome sequencing identifies association at ADCY7. *Nat Genet* 2017;49(2):186–92.
- [71] de Lange KM, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;49(2):256–61.
- [72] Halfvarson J. Genetics in twins with Crohn's disease: less pronounced than previously believed? *Inflamm Bowel Dis* 2011;17(1):6–12.
- [73] Manolio TA, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461(7265):747–53.
- [74] Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 2001;69(1):124–37.
- [75] Elding H, et al. Dissecting the genetics of complex inheritance: linkage disequilibrium mapping provides insight into Crohn disease. *Am J Hum Genet* 2011;89(6):798–805.
- [76] Zuk O, et al. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 2012;109(4):1193–8.
- [77] Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010;11(12):843–54.
- [78] Rivas MA, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet* 2011;43(11):1066–73.
- [79] Erickson AR, et al. Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One* 2012;7(11):e49138.
- [80] Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017;169(7):1177–86.
- [81] Knights D, et al. Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome Med* 2014;6(12):107.
- [82] Bonder MJ, et al. The effect of host genetics on the gut microbiome. *Nat Genet* 2016;48(11):1407–12.
- [83] Chu H, et al. Gene-microbiota interactions contribute to the pathogenesis of inflammatory bowel disease. *Science* 2016;352(6289):1116–20.
- [84] Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146(6):1489–99.
- [85] Gevers D, et al. The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15(3):382–92.
- [86] Waddington CH. The epigenotype. *Endeavour* 1942;1:42.
- [87] Ventham NT, et al. Beyond gene discovery in inflammatory bowel disease: the emerging role of epigenetics. *Gastroenterology* 2013;145(2):293–308.
- [88] Nimmo ER, et al. Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflamm Bowel Dis* 2012;18(5):889–99.
- [89] Cooke J, et al. Mucosal genome-wide methylation changes in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;18(11):2128–37.
- [90] Jenke AC, Zilbauer M. Epigenetics in inflammatory bowel disease. *Curr Opin Gastroenterol* 2012;28(6):577–84.
- [91] Mir SA, et al. Prenatal methyl-donor supplementation augments colitis in young adult mice. *PLoS One* 2013;8(8):e73162.
- [92] Schaible TD, et al. Maternal methyl-donor supplementation induces prolonged murine offspring colitis susceptibility in association with mucosal epigenetic and microbiomic changes. *Hum Mol Genet* 2011;20(9):1687–96.
- [93] Lister R, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 2009;462(7271):315–22.
- [94] Bock C, et al. Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol* 2010;28(10):1106–14.
- [95] Adams AT, et al. Two-stage genome-wide methylation profiling in childhood-onset Crohn's Disease implicates epigenetic alterations at the VMP1/MIR21 and HLA loci. *Inflamm Bowel Dis* 2014;20(10):1784–93.
- [96] Harris RA, et al. Genome-wide peripheral blood leukocyte DNA methylation microarrays identified a single association with inflammatory bowel diseases. *Inflamm Bowel Dis* 2012;18(12):2334–41.
- [97] Sadler T, et al. Cytokine-induced chromatin modifications of the type I collagen alpha 2 gene during intestinal endothelial-to-mesenchymal transition. *Inflamm Bowel Dis* 2013;19(7):1354–64.
- [98] Hasler R, et al. A functional methylome map of ulcerative colitis. *Genome Res* 2012;22(11):2130–7.
- [99] Weiser M, et al. Molecular classification of Crohn's disease reveals two clinically relevant subtypes. *Gut* 2016;67(1):36–42.
- [100] Bonder MJ, et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* 2017;49(1):131–8.



- [101] Barski A, et al. High-resolution profiling of histone methylations in the human genome. *Cell* 2007;129(4):823–37.
- [102] Wang Z, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 2008;40(7):897–903.
- [103] Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 2009;10(10):669–80.
- [104] Farnham PJ. Insights from genomic profiling of transcription factors. *Nat Rev Genet* 2009;10(9):605–16.
- [105] Mehta S, Jeffrey KL. Beyond receptors and signaling: epigenetic factors in the regulation of innate immunity. *Immunol Cell Biol* 2015;93(3):233–44.
- [106] Chi SW, et al. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 2009;460(7254):479–86.
- [107] Fairfax BP, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* 2014;343(6175):1246949.
- [108] Gibson G, Powell JE, Marigorta UM. Expression quantitative trait locus analysis for translational medicine. *Genome Med* 2015;7(1):60.
- [109] Vieira Braga FA, Teichmann SA, Chen X. Genetics and immunity in the era of single-cell genomics. *Hum Mol Genet* 2016;25(R2):R141–8.
- [110] Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* 2011;12(2):87–98.
- [111] Westermann AJ, Barquist L, Vogel J. Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathog* 2017;13(2):e1006033.
- [112] Planell N, et al. Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* 2013;62(7):967–76.
- [113] Rosen MJ, et al. Mucosal expression of type 2 and type 17 immune response genes distinguishes ulcerative colitis from colon-only Crohn's disease in treatment-naïve pediatric patients. *Gastroenterology* 2017;152(6):1345–57 e7.
- [114] Marigorta UM, et al. Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn's disease. *Nat Genet* 2017;49:1517–21.
- [115] Lin J, et al. Novel specific microRNA biomarkers in idiopathic inflammatory bowel disease unrelated to disease activity. *Mod Pathol* 2014;27(4):602–8.
- [116] Clark PM, et al. Bioinformatics analysis reveals transcriptome and microRNA signatures and drug repositioning targets for IBD and other autoimmune diseases. *Inflamm Bowel Dis* 2012;18(12):2315–33.
- [117] de Bruyn M, et al. Infliximab restores the dysfunctional matrix remodeling protein and growth factor gene expression in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2014;20(2):339–52.
- [118] Kalla R, et al. MicroRNAs: new players in IBD. *Gut* 2015;64(3):504–17.
- [119] Pandey A, Mann M. Proteomics to study genes and genomes. *Nature* 2000;405(6788):837–46.
- [120] Mann M, Jensen ON. Proteomic analysis of post-translational modifications. *Nat Biotechnol* 2003;21(3):255–61.
- [121] Bennike T, et al. Biomarkers in inflammatory bowel diseases: current status and proteomics identification strategies. *World J Gastroenterol* 2014;20(12):3231–44.
- [122] Barcelo-Batlioni S, et al. Proteomic analysis of cytokine induced proteins in human intestinal epithelial cells: implications for inflammatory bowel diseases. *Proteomics* 2002;2(5):551–60.
- [123] Fogt F, et al. Proteomic analysis of mucosal preparations from patients with ulcerative colitis. *Mol Med Rep* 2008;1(1):51–4.
- [124] Seeley EH, et al. Proteomic patterns of colonic mucosal tissues delineate Crohn's colitis and ulcerative colitis. *Proteomics Clin Appl* 2013;7(7-8):541–9.
- [125] Hisamatsu T, et al. Novel, objective, multivariate biomarkers composed of plasma amino acid profiles for the diagnosis and assessment of inflammatory bowel disease. *PLoS One* 2012;7(1):e31131.
- [126] Williams HR, et al. Serum metabolic profiling in inflammatory bowel disease. *Dig Dis Sci* 2012;57(8):2157–65.
- [127] Sharma U, et al. Similarity in the metabolic profile in macroscopically involved and un-involved colonic mucosa in patients with inflammatory bowel disease: an in vitro proton ((1)H) MR spectroscopy study. *Magn Reson Imaging* 2010;28(7):1022–9.
- [128] Schicho R, et al. Quantitative metabolomic profiling of serum, plasma, and urine by (1)H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. *J Proteome Res* 2012;11(6):3344–57.
- [129] Zamboni N, Saghatelian A, Patti GJ. Defining the metabolome: size, flux, and regulation. *Mol Cell* 2015;58(4):699–706.
- [130] Marchesi JR, et al. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 2007;6(2):546–51.
- [131] De Preter V, et al. Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in IBD. *Gut* 2015;64(3):447–58.
- [132] Hugot JP, et al. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;379(6568):821–3.
- [133] Hampe J, et al. A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. *Am J Hum Genet* 1999;64(3):808–16.
- [134] Cho JH, et al. Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: evidence for epistasis between 1p and IBD1. *Proc Natl Acad Sci USA* 1998;95(13):7502–7.
- [135] Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987;40(1):1–14.
- [136] Hugot JP, Zouali H, Lesage S. Lessons to be learned from the NOD2 gene in Crohn's disease. *Eur J Gastroenterol Hepatol* 2003;15(6):593–7.
- [137] Ogura Y, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;411(6837):603–6.
- [138] Hampe J, et al. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *Lancet* 2001;357(9272):1925–8.
- [139] Cummings JR, et al. The genetics of NOD-like receptors in Crohn's disease. *Tissue Antigens* 2010;76(1):48–56.
- [140] Inohara N, Nunez G. NODs: intracellular proteins involved in inflammation and apoptosis. *Nat Rev Immunol* 2003;3(5):371–82.
- [141] Inohara N, et al. Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease. *J Biol Chem* 2003;278(8):5509–12.
- [142] Harton JA, et al. Cutting edge: CATERPILLER: a large family of mammalian genes containing CARD, pyrin, nucleotide-binding, and leucine-rich repeat domains. *J Immunol* 2002;169(8):4088–93.



- [143] Kanazawa N, et al. Early-onset sarcoidosis and CARD15 mutations with constitutive nuclear factor-kappaB activation: common genetic etiology with Blau syndrome. *Blood* 2005;105(3):1195–7.
- [144] Miceli-Richard C, et al. CARD15 mutations in Blau syndrome. *Nat Genet* 2001;29(1):19–20.
- [145] Ho LP, et al. CARD 15 gene mutations in sarcoidosis. *Thorax* 2005;60(4):354–5.
- [146] Schurmann M, et al. CARD15 gene mutations in sarcoidosis. *Eur Respir J* 2003;22(5):748–54.
- [147] Chamaillard M, et al. An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat Immunol* 2003;4(7):702–7.
- [148] Girardin SE, et al. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003;278(11):8869–72.
- [149] Bonen DK, et al. Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan. *Gastroenterology* 2003;124(1):140–6.
- [150] Chamaillard M, et al. Nods, Nalps and Naip: intracellular regulators of bacterial-induced inflammation. *Cell Microbiol* 2003;5(9):581–92.
- [151] Inohara N, Ogura Y, Nunez G. Nods: a family of cytosolic proteins that regulate the host response to pathogens. *Curr Opin Microbiol* 2002;5(1):76–80.
- [152] Amendola A, et al. Nod2 deficiency is associated with an increased mucosal immunoregulatory response to commensal microorganisms. *Mucosal Immunol* 2014;7(2):391–404.
- [153] Strober W, Fuss I, Mannon P. The fundamental basis of inflammatory bowel disease. *J Clin Invest* 2007;117(3):514–21.
- [154] Watanabe T, et al. NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. *Nat Immunol* 2004;5(8):800–8.
- [155] Watanabe T, et al. Nucleotide binding oligomerization domain 2 deficiency leads to dysregulated TLR2 signaling and induction of antigen-specific colitis. *Immunity* 2006;25(3):473–85.
- [156] Hisamatsu T, et al. CARD15/NOD2 functions as an antibacterial factor in human intestinal epithelial cells. *Gastroenterology* 2003;124(4):993–1000.
- [157] Lala S, et al. Crohn's disease and the NOD2 gene: a role for Paneth cells. *Gastroenterology* 2003;125(1):47–57.
- [158] Wehkamp J, Stange EF. NOD2 mutation and mice: no Crohn's disease but many lessons to learn. *Trends Mol Med* 2005;11(7):307–9.
- [159] Wehkamp J, et al. NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. *Gut* 2004;53(11):1658–64.
- [160] Wehkamp J, et al. Defensin deficiency, intestinal microbes, and the clinical phenotypes of Crohn's disease. *J Leukoc Biol* 2005;77(4):460–5.
- [161] Wehkamp J, et al. Mechanisms of disease: defensins in gastrointestinal diseases. *Nat Clin Pract Gastroenterol Hepatol* 2005;2(9):406–15.
- [162] Wehkamp J, et al. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proc Natl Acad Sci USA* 2005;102(50):18129–34.
- [163] Wehkamp J, Fellermann K, Stange EF. Human defensins in Crohn's disease. *Chem Immunol Allergy* 2005;86:42–54.
- [164] Kobayashi KS, et al. Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* 2005;307(5710):731–4.
- [165] Maeda S, et al. Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing. *Science* 2005;307(5710):734–8.
- [166] Mondot S, et al. Altered gut microbiota composition in immune-impaired Nod2(-/-) mice. *Gut* 2012;61(4):634–5.
- [167] Petnicki-Ocwieja T, et al. Nod2 is required for the regulation of commensal microbiota in the intestine. *Proc Natl Acad Sci USA* 2009;106(37):15813–18.
- [168] Rehman A, et al. Nod2 is essential for temporal development of intestinal microbial communities. *Gut* 2011;60(10):1354–62.
- [169] Abreu MT, et al. Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. *Gastroenterology* 2002;123(3):679–88.
- [170] Ahmad T, et al. The molecular classification of the clinical manifestations of Crohn's disease. *Gastroenterology* 2002;122(4):854–66.
- [171] Cuthbert AP, et al. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology* 2002;122(4):867–74.
- [172] Brant SR, Shugart YY. Inflammatory bowel disease gene hunting by linkage analysis: rationale, methodology, and present status of the field. *Inflamm Bowel Dis* 2004;10(3):300–11.
- [173] Linde K, et al. Card15 and Crohn's disease: healthy homozygous carriers of the 3020insC frameshift mutation. *Am J Gastroenterol* 2003;98(3):613–17.
- [174] Radlmayr M, et al. The c-insertion mutation of the NOD2 gene is associated with fistulizing and fibrostenotic phenotypes in Crohn's disease. *Gastroenterology* 2002;122(7):2091–2.
- [175] Economou M, et al. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a meta-analysis. *Am J Gastroenterol* 2004;99(12):2393–404.
- [176] Vind I, et al. NOD2/CARD15 gene polymorphisms in Crohn's disease: a genotype-phenotype analysis in Danish and Portuguese patients and controls. *Digestion* 2005;72(2–3):156–63.
- [177] Arnott ID, et al. NOD2/CARD15, TLR4 and CD14 mutations in Scottish and Irish Crohn's disease patients: evidence for genetic heterogeneity within Europe? *Genes Immun* 2004;5(5):417–25.
- [178] Russell RK, et al. Genotype-phenotype analysis in childhood-onset Crohn's disease: NOD2/CARD15 variants consistently predict phenotypic characteristics of severe disease. *Inflamm Bowel Dis* 2005;11(11):955–64.
- [179] Bairead E, et al. Association of NOD2 with Crohn's disease in a homogenous Irish population. *Eur J Hum Genet* 2003;11(3):237–44.
- [180] Helio T, et al. CARD15/NOD2 gene variants are associated with familiarly occurring and complicated forms of Crohn's disease. *Gut* 2003;52(4):558–62.
- [181] Kugathasan S, et al. Comparative phenotypic and CARD15 mutational analysis among African American, Hispanic, and White children with Crohn's disease. *Inflamm Bowel Dis* 2005;11(7):631–8.

- [182] Lee GH, et al. [Frequency analysis of NOD2 gene mutations in Korean patients with Crohn's disease]. *Korean J Gastroenterol* 2005;45(3):162–8.
- [183] Croucher PJ, et al. Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations. *Eur J Hum Genet* 2003;11(1):6–16.
- [184] Inoue N, et al. Lack of common NOD2 variants in Japanese patients with Crohn's disease. *Gastroenterology* 2002;123(1):86–91.
- [185] Yamazaki K, et al. Absence of mutation in the NOD2/CARD15 gene among 483 Japanese patients with Crohn's disease. *J Hum Genet* 2002;47(9):469–72.
- [186] Leong RW, et al. NOD2/CARD15 gene polymorphisms and Crohn's disease in the Chinese population. *Aliment Pharmacol Ther* 2003;17(12):1465–70.
- [187] Guo QS, et al. NOD2 3020insC frameshift mutation is not associated with inflammatory bowel disease in Chinese patients of Han nationality. *World J Gastroenterol* 2004;10(7):1069–71.
- [188] Zhou Z, et al. Variation at NOD2/CARD15 in familial and sporadic cases of Crohn's disease in the Ashkenazi Jewish population. *Am J Gastroenterol* 2002;97(12):3095–101.
- [189] Lanzarotto F, Akbar A, Ghosh S. Does innate immune response defect underlie inflammatory bowel disease in the Asian population? *Postgrad Med J* 2005;81(958):483–5.
- [190] Cleynen I, et al. Genetic factors conferring an increased susceptibility to develop Crohn's disease also influence disease phenotype: results from the IBDchip European Project. *Gut* 2013;62:1556–65.
- [191] Lesage S, et al. CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. *Am J Hum Genet* 2002;70(4):845–57.
- [192] Cho JH. The Nod2 gene in Crohn's disease: implications for future research into the genetics and immunology of Crohn's disease. *Inflamm Bowel Dis* 2001;7(3):271–5.
- [193] Schreiber S, et al. Genetics of Crohn disease, an archetypal inflammatory barrier disease. *Nat Rev Genet* 2005;6(5):376–88.
- [194] Pathan S, et al. Deep dermatophytosis and inherited CARD9 deficiency. *N Engl J Med* 2013;369(18):1704–14.
- [195] Glocker EO, et al. A homozygous CARD9 mutation in a family with susceptibility to fungal infections. *N Engl J Med* 2009;361(18):1727–35.
- [196] Hsu YM, et al. The adaptor protein CARD9 is required for innate immune responses to intracellular pathogens. *Nat Immunol* 2007;8(2):198–205.
- [197] Cagliani R, et al. Balancing selection is common in the extended MHC region but most alleles with opposite risk profile for autoimmune diseases are neutrally evolving. *BMC Evol Biol* 2011;11:171.
- [198] Neefjes J, et al. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol* 2011;11(12):823–36.
- [199] Yap LM, Ahmad T, Jewell DP. The contribution of HLA genes to IBD susceptibility and phenotype. *Best Pract Res Clin Gastroenterol* 2004;18(3):577–96.
- [200] Ahmad T, Marshall SE, Jewell D. Genetics of IBD: Topic highlights for the role of the HLA complex. *World J Gastroenterol* 2006;12(23):3628–35.
- [201] Ahmad T, Marshall SE, Jewell D. Genetics of inflammatory bowel disease: the role of the HLA complex. *World J Gastroenterol* 2006;12(23):3628–35.
- [202] Stokkers PC, et al. HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis. *Gut* 1999;45(3):395–401.
- [203] Hampe J, et al. Linkage of inflammatory bowel disease to human chromosome 6p. *Am J Hum Genet* 1999;65(6):1647–55.
- [204] Trachtenberg EA, et al. HLA class II haplotype associations with inflammatory bowel disease in Jewish (Ashkenazi) and non-Jewish Caucasian populations. *Hum Immunol* 2000;61(3):326–33.
- [205] Yoshitake S, et al. HLA class II alleles in Japanese patients with inflammatory bowel disease. *Tissue Antigens* 1999;53(4 Pt 1):350–8.
- [206] Ahmad T, et al. The contribution of human leucocyte antigen complex genes to disease phenotype in ulcerative colitis. *Tissue Antigens* 2003;62(6):527–35.
- [207] Futami S, et al. HLA-DRB1\*1502 allele, subtype of DR15, is associated with susceptibility to ulcerative colitis and its progression. *Dig Dis Sci* 1995;40(4):814–18.
- [208] Roussomoustakaki M, et al. Genetic markers may predict disease behavior in patients with ulcerative colitis. *Gastroenterology* 1997;112(6):1845–53.
- [209] Bouma G, et al. Evidence for genetic heterogeneity in inflammatory bowel disease (IBD): HLA genes in the predisposition to suffer from ulcerative colitis (UC) and Crohn's disease (CD). *Clin Exp Immunol* 1997;109(1):175–9.
- [210] Hancock L, et al. Clinical and molecular characteristics of isolated colonic Crohn's disease. *Gut* 2006;55(Suppl II):A1.
- [211] Silverberg MS, et al. A population- and family-based study of Canadian families reveals association of HLA DRB1\*0103 with colonic involvement in inflammatory bowel disease. *Inflamm Bowel Dis* 2003;9(1):1–9.
- [212] Ahmad T, et al. Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. *Hum Mol Genet* 2003;12(6):647–56.
- [213] Akolkar PN, et al. The IBD1 locus for susceptibility to Crohn's disease has a greater impact in Ashkenazi Jews with early onset disease. *Am J Gastroenterol* 2001;96(4):1127–32.
- [214] Ahmad T, et al. High resolution MIC genotyping: design and application to the investigation of inflammatory bowel disease susceptibility. *Tissue Antigens* 2002;60(2):164–79.
- [215] van Heel DA, et al. Inflammatory bowel disease is associated with a TNF polymorphism that affects an interaction between the OCT1 and NF-(kappa)B transcription factors. *Hum Mol Genet* 2002;11(11):1281–9.
- [216] O'Callaghan NJ, et al. Association of TNF-alpha-857C with inflammatory bowel disease in the Australian population. *Scand J Gastroenterol* 2003;38(5):533–4.
- [217] Orchard TR, et al. Clinical phenotype is related to HLA genotype in the peripheral arthropathies of inflammatory bowel disease. *Gastroenterology* 2000;118(2):274–8.

- [218] Pathan S, et al. Confirmation of the novel association at the BTNL2 locus with ulcerative colitis. *Tissue Antigens* 2009;74(4):322–9.
- [219] Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet* 2011;12(10):683–91.
- [220] Maurano MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 2012;337(6099):1190–5.
- [221] Buenrostro JD, et al. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* 2015;109:21.29.1–21.29.9.
- [222] Moffatt MF, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;448(7152):470–3.
- [223] Goyette P, et al. Gene-centric association mapping of chromosome 3p implicates MST1 in IBD pathogenesis. *Mucosal Immunol* 2008;1(2):131–8.
- [224] Pathan S, et al. Putative genetic loci and confirmation of association in IBD. *Gastroenterology* 2010;138(5):S-6.
- [225] Satsangi J, et al. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7, and 12. *Nat Genet* 1996;14(2):199–202.
- [226] Travassos LH, et al. Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* 2010;11(1):55–62.
- [227] Cooney R, et al. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med* 2010;16(1):90–7.
- [228] Brain O, et al. Functional consequences of mutations in the autophagy genes in the pathogenesis of Crohn's disease. *Inflamm Bowel Dis* 2012;18(4):778–881.
- [229] Cadwell K, et al. A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. *Nature* 2008;456(7219):259–63.
- [230] Brest P, et al. A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat Genet* 2011;43(3):242–5.
- [231] Brain O, et al. The intracellular sensor NOD2 induces microRNA-29 expression in human dendritic cells to limit IL-23 release. *Immunity* 2013;39(3):521–36.
- [232] Glocker EO, et al. IL-10 and IL-10 receptor defects in humans. *Ann N Y Acad Sci* 2011;1246:102–7.
- [233] Consortium, UK IBD Genetics, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet* 2009;41(12):1330–4.
- [234] Asano K, et al. A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. *Nat Genet* 2009;41(12):1325–9.
- [235] Ahn SH, et al. Hepatocyte nuclear factor 4alpha in the intestinal epithelial cells protects against inflammatory bowel disease. *Inflamm Bowel Dis* 2008;14(7):908–20.
- [236] Fisher SA, et al. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 2008;40(6):710–12.
- [237] Boltin D, et al. Mucin function in inflammatory bowel disease: an update. *J Clin Gastroenterol* 2013;47(2):106–11.
- [238] Johansson ME, et al. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci USA* 2008;105(39):15064–9.
- [239] Kashyap PC, et al. Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. *Proc Natl Acad Sci USA* 2013;110(42):17059–64.
- [240] Tong M, et al. Reprogramming of gut microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. *ISME J* 2014;8(11):2193–206.
- [241] Frank DN, et al. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci USA* 2007;104(34):13780–5.
- [242] Lee JC, et al. Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. *J Clin Invest* 2011;121(10):4170–9.
- [243] Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet* 2009;10(1):43–55.
- [244] Yamagata K, et al. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 1996;384(6608):458–60.
- [245] Kathiresan S, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009;41(1):56–65.
- [246] Snook JA, de Silva HJ, Jewell DP. The association of autoimmune disorders with inflammatory bowel disease. *Q J Med* 1989;72(269):835–40.
- [247] Murdoch TB, et al. Prevalence of genetic variants associated with inflammatory bowel disease in a healthy First Nations cohort. *CMAJ* 2012;184(8):E435–41.
- [248] Orstavik KH, et al. Novel splicing mutation in the NEMO (IKK-gamma) gene with severe immunodeficiency and heterogeneity of X-chromosome inactivation. *Am J Med Genet A* 2006;140(1):31–9.
- [249] Puel A, et al. The NEMO mutation creating the most-upstream premature stop codon is hypomorphic because of a reinitiation of translation. *Am J Hum Genet* 2006;78(4):691–701.
- [250] Hayward PA, Satsangi J, Jewell DP. Inflammatory bowel disease and the X chromosome. *Q J Med* 1996;89(9):713–18.
- [251] Glocker E, Grimbacher B. Inflammatory bowel disease: is it a primary immunodeficiency? *Cell Mol Life Sci* 2012;69(1):41–8.
- [252] Valentonyte R, et al. Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat Genet* 2005;37(4):357–64.
- [253] Hirschfield GM, et al. Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med* 2009;360(24):2544–55.
- [254] Roda G, et al. New proteomic approaches for biomarker discovery in inflammatory bowel disease. *Inflamm Bowel Dis* 2010;16(7):1239–46.

- [255] Lee JC, et al. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat Genet* 2017;49(2):262–8.
- [256] Heap GA, et al. HLA-DQA1-HLA-DRB1 variants confer susceptibility to pancreatitis induced by thiopurine immunosuppressants. *Nat Genet* 2014;46(10):1131–4.
- [257] Yang SK, et al. A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. *Nat Genet* 2014;46(9):1017–20.
- [258] Sandborn WJ, et al. Vedolizumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2013;369(8):711–21.
- [259] Feagan BG, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013;369(8):699–710.
- [260] Sivan A, et al. Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* 2015;350(6264):1084–9.
- [261] Vetizou M, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* 2015;350(6264):1079–84.
- [262] Zhang W, Li F, Nie L. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. *Microbiology* 2010;156(Pt 2):287–301.
- [263] Klionsky DJ. Crohn's disease, autophagy, and the Paneth cell. *N Engl J Med* 2009;360(17):1785–6.
- [264] Coskun M, et al. MicroRNAs in inflammatory bowel disease—pathogenesis, diagnostics and therapeutics. *World J Gastroenterol* 2012;18(34):4629–34.
- [265] Palmieri O, et al. Functional implications of MicroRNAs in Crohn's disease revealed by integrating MicroRNA and messenger RNA expression profiling. *Int J Mol Sci* 2017;18:7.
- [266] Beres NJ, et al. Altered mucosal expression of microRNAs in pediatric patients with inflammatory bowel disease. *Dig Liver Dis* 2017;49(4):378–87.
- [267] Navin N, Hicks J. Future medical applications of single-cell sequencing in cancer. *Genome Med* 2011;3(5):31.

## Further reading

Altmuller J, et al. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 2001;69(5):936–50.

Davies JL, et al. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 1994;371(6493):130–6.

Lees CW, et al. New IBD genetics: common pathways with other diseases. *Gut* 2011;60(12):1739–53.

# Molecular biology of acute and chronic inflammation

Dhavendra Kumar<sup>1,2,3</sup>

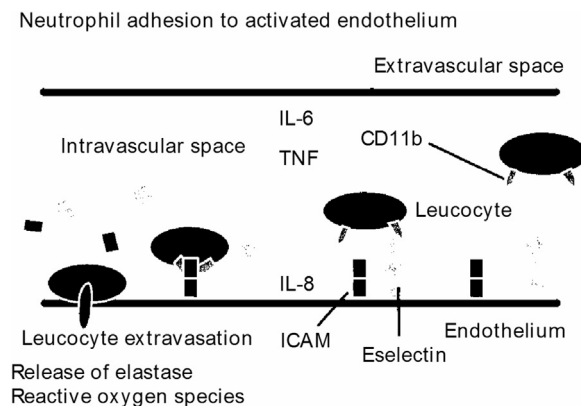
<sup>1</sup>Institute of Medical Genetics, University Hospital of Wales, Cardiff, United Kingdom <sup>2</sup>Genomic Policy Unit, University of South Wales, Pontypridd, United Kingdom <sup>3</sup>William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

## 22.1 Introduction

Inflammation is a natural response of tissues and organs following acute or chronic trauma, acute infection, or acute immune-mediated tissue damage. It is spontaneous and often positive and beneficial in response to acute injury or invading pathogen.

New advances in molecular medicine have enhanced understanding of inflammatory response to injury and shock at the molecular level. Several mechanisms are under consideration for the development of acute inflammation following acute trauma or sepsis. Among these main molecular mechanisms include the macrophage theory, the gut hypothesis, the two-hit theory, and the microvascular environmental theory (Fig. 22.1) [1,2].

Lately, the neutrophil-mediated tissue-injury theory has gained much acceptance. It suggests adherence of activated neutrophils to the endothelium resulting in creation of local intercellular microenvironment, isolated from circulating oxidant scavengers and antiproteases. This mechanism results in very high concentrations of toxic metabolites resulting in unopposed endothelial injury. A number of molecular markers are now available, which can predict severity of endothelial injury and offer opportunity for targeted therapeutic intervention and may



**FIGURE 22.1** Molecular display of the microvascular environment theory [1]. Source: Adapted with permission from Giannoudis PV, Hildebrand F, Pape HC. Inflammatory serum markers in patients with multiple trauma—can they predict outcome? *J Bone Joint Surgery (Br)* 2004;86-B (3):313–23.



prevent the onset of adverse sequelae. The current knowledge on the clinical effectiveness of existing inflammatory markers of immune reactivity is evaluated in this chapter.

## 22.2 Molecular pathology of acute inflammation (sepsis and trauma)

Inflammation is a spontaneous response to traumatic tissue injury or an invading pathogen complicating with sepsis. It can be a beneficial acute, transient immune response that facilitates tissue repair, turnover, and adaptation of many tissues. However, the acute inflammatory response to pathogen-associated molecular mechanisms may be impaired during aging, resulting in increased susceptibility to infection. In contrast, chronic inflammation is usually of low grade and persistent, resulting in complex tissue responses that lead to tissue degeneration.

A number of regulatory proteins produced and secreted by lymphocytes and other cells have a role in cascading the acute inflammatory response to injury and infection. Following acute challenges, such as trauma, patients are subjected to dynamic alterations in the hemodynamic, metabolic, and immune responses that are largely orchestrated by endogenous mediators referred to as cytokines. In an acute inflammatory state a patient is loaded with pulsating cytokines [3]. This state could have fatal consequences and is part of the normal response to the systemic inflammatory response syndrome (SIRS) followed by a period mediated by compensatory antiinflammatory syndrome (CARS) [4]. While these initial acute inflammatory responses could have beneficial effects, they may lead to adult respiratory distress syndrome (ARDS) or multiple tissue/organ damage similar to multiple organ dysfunction syndrome (MODS) if they remain progressive or untreated [5].

Experts on acute versus chronic inflammation discuss “one-hit” and “two-hit” models. According to “one-hit” model, the initial massive injury and shock give rise to an intense systemic inflammation that causes activation of the innate immune system, including macrophages, leukocytes, natural killer cells, and inflammatory cell migration, enhanced by the production of interleukin (IL)-8 and complement components (C5a and C3a) [6]. In the “two-hit” model the stimulus is less intense and normally resolved. However, the patient may remain vulnerable to secondary inflammatory challenge and may respond similar to ARDS or MODS [7]. Hyperstimulation by single or multiple hits is considered by many to be the key element in the pathogenesis of ARDS or MODS. The second or repeat insults may take many forms, commonly acute infections, or surgical operations. The clinician, faced with this situation, will need to make difficult decisions keeping best interests of the patient.

A number of different inflammatory mediators are now known based on studies on patients subjected to acute trauma or sepsis [8]. The response initially corresponds to the first-hit theory and is subsequently triggered by the second-hit, commonly a surgical procedure. The models of “one-hit” and “second-hit” are now the basis of management plans in any acute or critical care unit. At the molecular level a variety of inflammatory mediators are implicated in single or multiple organ dysfunction. These are grouped as acute-phase reactants, mediator activity, and cellular activity (Table 22.1) [3].

In clinical practice, for a marker to be useful, it must be both sensitive and provide best possible specific information for severity and extent of local or systemic inflammation. The marker should be easily and reliably measured from a blood sample. An ideal marker will accurately discriminate between those with a given condition and those without. The clinical utility of the marker is determined by its sensitivity to distinguish between the inflammatory response and infection allowing earlier identification of patients with a specific disease process. In this context, C-reactive protein (CRP) is probably the most widely used marker in clinical practice. However, it does not have sufficient discriminatory power from current and previous injury. Further, it is not a good marker to assess the long-term outcome. Several studies on cytokines and ILs have led to selection of two markers in

**TABLE 22.1** Major groups of the serum inflammatory markers [3].

Group	Serum inflammatory markers
Acute-phase reactants	LBP, CRP, PCT
Mediator activity	TNF, IL-1, IL-6, IL-10, IL-18
Cellular activity	TNF-RI, TNF-RII, IL-1R, sIL-6-R, mIL-6- $\alpha$ , ICAM-1 Eselectin, CD11b, elastase, HLA-DR class-II antigens, DNA

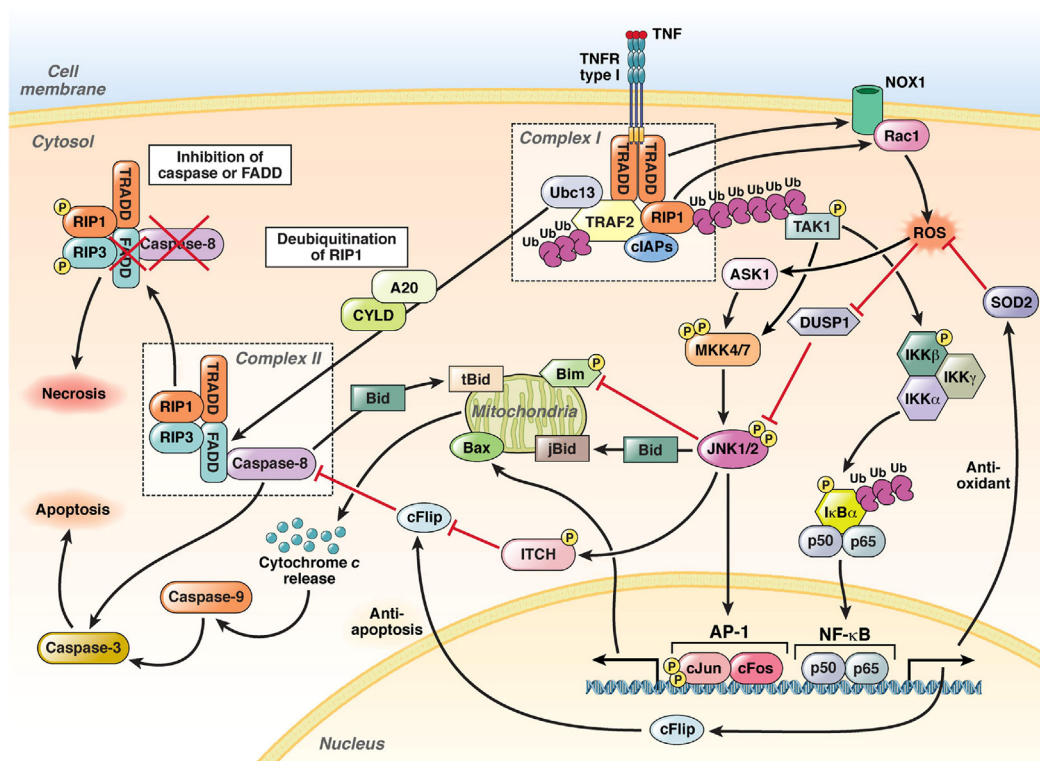
CRP, C-reactive protein; HLA, human leukocyte antigen; ICAM, intercellular adhesion molecule; IL, interleukin; LBP, lipopolysaccharide-binding protein; PCT, procalcitonin; TNF, tumor necrosis factor.

clinical practice—IL-6 for systemic low-grade inflammation and LBP for acute and early marker of inflammation. The use of these two markers together may offer ability to detect the onset of SIRS and allow early intervention to prevent MODS, to distinguish between inflammation and infection and to monitor the response to standard and innovative therapies [9].

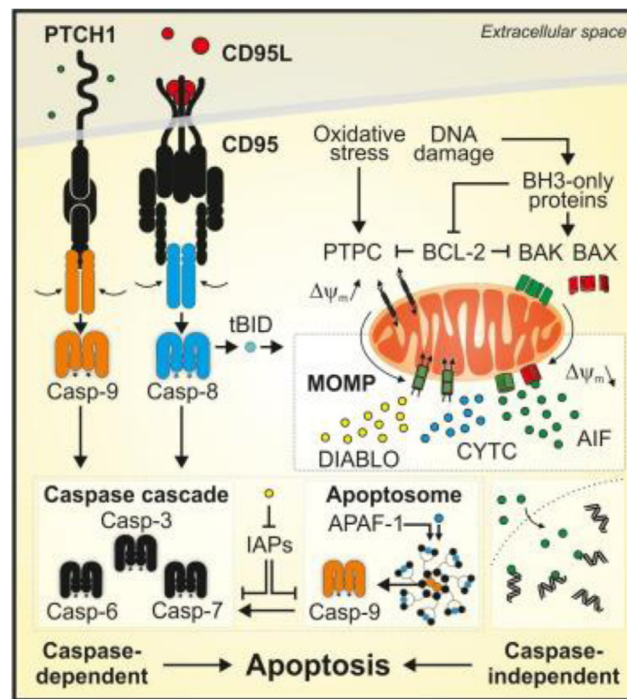
## 22.3 Molecular pathology of chronic inflammation

Chronic inflammation has many features of acute inflammation but is usually of low grade and persistent, resulting in pathological responses that lead to tissue degeneration, and may trigger carcinogenesis. At the molecular level, both acute and chronic inflammation share similar mechanisms that constitute complex inflammatory pathways mediated by several genes and molecules in the extracellular, cell membrane, intracellular domains (Fig. 22.2). In most cases, sequential and recurrent cellular and tissue damage leads to fibrosis, for example, liver fibrosis (cirrhosis) due to chronic alcohol-induced inflammation or obesity associated with overnutrition [11]. Alternatively, the chronic inflammatory processes may also progress to cancer, for example, prostate cancer that follows long-standing chronic prostatitis with intermittent episodes of acute prostatitis [12] and similarly hepatic carcinoma (Fig. 22.2) [10].

Several biological studies provide evidence that active oxygen and nitrogen radicals produced by chronic inflammatory tissue increased the risk of cancer by suppressing antitumor activity and stimulating carcinogenesis [13].



**FIGURE 22.2** JNKs in TNF signaling. (1) Binding of TNF to the TNFR type I leads to the rapid formation of complex I, comprising TRADD, RIP1, TRAF2, cIAP1, cIAP2, and Ubc13. cIAP-mediated K63 ubiquitination of RIP1 recruits and activates TAK1. (2) MAP3Ks (TAK1 and ASK1) activate JNK1 and 2 through MKK4 and 7. JNK activates AP-1, which comprises c-Jun and c-Fos. Simultaneously, JNK1 phosphorylates ITCH to ubiquitinate c-FLIP, which promotes caspase-8-dependent apoptosis. JNKs can also induce mitochondria-dependent apoptosis through Bax and degradation of Bim. (3) TAK1 phosphorylates and activates the IKK complex, which leads to phosphorylation, ubiquitination, and degradation of IκBα, resulting in nuclear translocation and activation of NF-κB, which comprises the p50 and p65 subunits. NF-κB induces the transcription of superoxide dismutase 2 (SOD2) and c-FLIP to prevent ROS production and caspase-8 activation, respectively. (4) Complex I also contributes to ROS production through reduced nicotinamide adenine dinucleotide phosphate oxidase 1 (NOX1) and Rac1. (5) Following formation of complex I, RIP1 is deubiquitinated by cylindromatosis (CYLD) or A20 to form complex II, comprising TRADD, FADD, RIP1, RIP3, and caspase-8. Normally, caspase-8 induces apoptosis. (6) However, if caspase-8 or FADD is blocked, RIP1 and RIP3 are phosphorylated and cause necrosis. Source: Adapted with permission from Seki E, Brenner DA, Karin M. A liver full of JNK: signaling in regulation of cell function and disease pathogenesis, and clinical approaches. *Gastroenterology* 2012;143(2):307–20.



**FIGURE 22.3** Molecular mechanisms of apoptosis. Extrinsic apoptosis can lead to either ligation of death receptors (e.g., CD95) or drop in the concentration of dependence receptor (e.g., PTCH1) ligands below a specific threshold. In particular, CD95 ligand (CD95L) stabilizes CD95 trimers, hence allowing for the assembly of a plasma membrane-associated supramolecular complex that favors the proximity-induced activation of caspase-8. In turn, active caspase-8 sets off the executioner machinery of apoptosis by activating a proteolytic cascade involving caspase-3, -6, and -7. Conversely, PTCH1 appears to mediate the activation of executioner caspases via caspase-9-dependent signal transduction cascade. Several intracellular stress conditions (e.g., DNA damage) are specifically sensed by small members of the BCL-2 protein family (BH3-only proteins), which activate mitochondrial outer membrane permeabilization (MOMP) by stimulating the pore-forming activity of BAX and BAK. Alternatively, MOMP can be initiated at the inner mitochondrial membrane by the unspecific opening of the “permeability transition pore complex” (PTPC). Both these lethal cascades can be held in check by antiapoptotic BCL-2 proteins, which physically bind (hence inhibiting) not only their pro-apoptotic counterparts but also various PTPC components. MOMP is paralleled by the dissipation of the mitochondrial transmembrane potential ( $\Delta\Psi_m$ ) and results in the liberation of several mitochondrial proteins such as cytochrome C (CYTC), which, together with the cytosolic adaptor APAF-1, generates the caspase-9-activating platform known as “apoptosome”; apoptosis-inducing factor (AIF), which exerts caspase-independent pro-apoptotic functions by mediating large-scale DNA fragmentation; and direct IAP-binding protein with low pI (DIABLO), which physically antagonizes various inhibitor of apoptosis proteins (IAPs). Thus, MOMP activates both caspase-dependent and caspase-independent mechanisms of apoptosis. Of note, the caspase-8-mediated cleavage of the BH3-only protein BID constitutes a major link between the extrinsic and intrinsic apoptotic pathways. *tBID*, Truncated BID. Source: Adapted with permission from Brenner C, Galluzzi L, Kepp O, Kroemer G. Decoding cell death signals in liver inflammation. *J Hepatol* 2013;59(3):583–94 [16].

New genetic evidence suggests that transcription factors, NF- $\kappa$ B and STAT3, play a role in the association between inflammation and cancer (Fig. 22.3) [14].

There are several possible mechanisms of chronic inflammation [15]:

- Persistent production of reactive molecules by infiltrating leukocytes that eventually damages the structural and cellular elements of tissues.
- Damaged nonimmune cells and activated immune cells, which result in the production of cytokines and amplify or modulate the inflammatory response disrupting normal cellular function.
- Interference with “anabolic signaling” by downregulating insulin through multiple and complex interactions of IL-6, tumor necrosis factor alpha (TNF- $\alpha$ ), insulin-like growth factor-1, erythropoietin signaling, and enhanced protein synthesis after a meal or bout of exercise.

## 22.4 Age-associated chronic inflammation

Aging is a natural physiological process that results from environmental, opportunistic, genetic, and epigenetic events in different cell types and tissues. These events are interactive and occur practically throughout life. An

important and common feature of aging tissues and organs is the chronic inflammation, often referred to as “inflammaging,” and described as low-grade chronic and systemic inflammation in the absence of overt infection or so-called sterile inflammation [17]. It is a highly significant risk factor for both morbidity and mortality in the elderly people. In most cases, there is a state of mild inflammation revealed by elevated few biomarkers such as CRP and IL-6. These biomarkers can predict many aging phenotypes, such as changes in body composition, energy production and utilization, metabolic homeostasis, immune senescence, and neuronal health. Precise underlying mechanisms for “inflammaging” remain unknown. However, it is important to identify pathways that govern age-related inflammation across multiple systems. Many aged tissues probably remain in chronically inflamed state, however without any signs of infection. These basic facts are important for developing management and treatment strategies for chronic inflammation in elderly people.

Inflammaging is a complex process that involves a number of genetic and environmental factors. There are likely multiple sources [15]:

1. Accumulation of damaged macromolecules and cells (self-debris) due to increased production and/or inadequate elimination. Self-debris can mimic bacterial agents and function as “endogenous damage” that activates innate immunity. A network of inflammatory sensors, including the Nlrp3 inflammasome, recognizes the endogenous damage signals and initiates immune reactions that are necessary for physiological repair. The endogenous damage signals include damaged cellular and organelle components, free radicals from oxidative stress, metabolites, such as extracellular adenosine tri phosphate (ATP), fatty acids, urate crystals, ceramides, cardiolipin, amyloid, succinate, peroxidized lipids, advanced glycogen end products, altered *N*-glycans, and NMGB1.
2. Release of harmful products from the normal microbial constituents of the human body such as oral and gut microbiomes. It is likely that due to aging process itself, the gut itself is unable to sequester microbes, and thus a state of low-grade inflammation sets in. Alternatively, the nature of gut microbial composition could also alter by age and elicit low-level gastrointestinal tract inflammation. These inflammatory by-products could then leak into the surrounding tissues and eventually into the circulation. Other pathological consequences of the low-grade gut inflammation could be disruption of the host–pathogen balance resulting in opportunistic upsurge of other pathogens such as cytomegalovirus (CMV) and Epstein–Barr viruses. Further, alterations in the “immune risk profile” due to excess of cytokines, notably IL-6 and TNF- $\alpha$  R1, could lead to enhanced risks of mortality and morbidity.

The pathogen-induced activation of endogenous proinflammatory molecular pattern changes includes activation of Nlrp3 inflammasome, secretion of cytokines IL-1 $\beta$ , IL-8, and cardiolipin. This in turn also leads to mitochondrial dysfunction that is also capable of activating Nlrp3 inflammasome. The net outcome of all these complex molecular interactions is the persistent state of low-grade inflammation in the elderly (inflammaging).

3. Chronic low-grade inflammation in the elderly is to some extent directly related to cellular senescence. Senescence is a cellular response to damage and stress. The physiological senescence response prevents cancer by suppressing the proliferation of cells with a compromised genome and contributes to optimal wound healing in normal tissues. However, conversely persistent senescence could drive aging and age-associated chronic morbid state due to excess secretion of proinflammatory cytokines, referred to as “senescence-associated secretory phenotype” or SASP that modifies the tissue microenvironment of target and adjoining normal tissues. Senescent cells accumulate with age in many tissues, notably the adipose visceral tissue, and are related to many age-related disorders [18].
4. Increased inflammation in the elderly may derive from enhanced activation of the coagulation system with age. Coagulation may be considered part of the inflammation system with many shared components and strong interactions. The increased hypercoagulable state noted in the elderly is probably a direct consequence of aging and could manifest with increased incidence of arterial and venous thrombosis; examples include endotoxemic shock due to microbial translocation, atherosclerotic plaque erosion, or loss of structural vascular integrity leading to stasis.
5. Inflammaging in the elderly is probably related to age-related immune changes or immunosenescence. Innate immunity in the elderly is maintained or even slightly enhanced compared to the adaptive immunity that declines with age. The age-related changes in the immune profile probably result from lifelong exposures to pathogen and antigens, as well as intrinsic changes in immune cells. Genetic predisposition is also likely to play an important role. Accelerated immunosenescence with aging is associated with certain viral infections such as CMV and HIV [19].



6. Defective or inappropriate regulation of the complement pathway can lead to local inflammatory reactions in the elderly. A notable example is age-related macular degeneration—the leading cause of blindness in the elderly people. This phenomenon is likely to be causally related to many other degenerative diseases in the elderly [20].

In clinical settings, molecular profiling of inflammaging is important for developing preventive and therapeutic strategies. There are many innovative modalities under development. Agents that eliminate senescent cells, or suppress their SASP, hold promise for diminishing chronic inflammation in the elderly. In a mouse model, reduction in the Nlrp3 inflammasome-dependent proinflammatory cascade attenuated age-related degenerative changes across multiple organs [21]. This is aided by exploration of new targets in signaling and transcriptional pathways that drive chronic inflammation. These new targets could become the focus of therapeutic interventions for managing low-grade inflammation in the elderly. Since immunosenescence is important, strategies such as thymic replacement, local or systemic neutralization of cytokines, such as IL-6, or development of specific vaccines against CMV and HIV may be important for stimulating adaptive immunity, and this minimizing impact of immune–system driven chronic inflammation in the elderly. Other strategies promoted include the use of pro and prebiotics and fecal microbiota transplantation to reduce the burden of inflammaging and age-related pathologies. Not surprisingly, the public health promoters are actively pursuing these interventions at the population level.

## 22.5 Molecular diagnosis and treatment of chronic inflammatory diseases

Chronic inflammatory diseases are ubiquitous involving clinically every body system, organ, and localized tissues. Essentially, there are four triggering mechanistic factors: (1) persistent or recurrent trauma (Raynaud's disease), (2) chronic and recurrent infection (mycobacterial diseases), (3) autoimmune-induced chronic inflammation [systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), chronic pancreatitis, ulcerative colitis, multiple sclerosis, etc.], and (4) persistent and recurrent vascular injury (atherosclerosis). In addition, other factors include noninfective and nonspecific chronic inflammation (sarcoidosis and amyloidosis), multifactorial chronic airway inflammation (bronchial asthma), and multifactorial polygenic diseases with chronic inflammation (age-related macular degeneration, bronchial asthma, psoriasis, and others). Some of these conditions are covered in other chapters and thus not discussed here.

### 22.5.1 Genomic and molecular diagnosis

In most specific acute and chronic inflammatory diseases, the evidence-based clinical diagnosis is based on the recent and past medical history, clinical symptoms and signs, and relevant imaging signs (conventional radiography, ultrasound scan, CT scan, and selective or whole-body MRI scan). It is routine practice to arrange standard hematological and biochemical (serum) markers in any patient presenting with acute inflammation, for example, acute trauma or sepsis. There are now evidence-based investigation guidelines commonly used in most hospitals in many different countries, which are often referred to as “bundles”, part of the clinical protocol and pathway recommended by the relevant professional organization, for example in the United Kingdom, the National Institute of Clinical Excellence (NICE) or the Royal College of Emergency Medicine have produced professionally validated guidance available on the public domain. The use of “sepsis bundle” is now universally used across the National Health Service (NHS) across all UK hospitals. This approach is hugely beneficial for instituting the prompt antimicrobial administration. Standard battery of biochemical serum markers, for example, CRP, may allow assessment of the state of acute inflammation induced either by an invading pathogen or trauma. However, this is nonspecific, and normal or low levels might not exclude the acute inflammation. The use of genetic or specific molecular diagnosis in acute inflammation is very limited. However, newer genome-based microbial infection diagnostic kit (e.g., *Oxford Nanopore*) might be useful in early diagnosis of specific pathogen involved in lower respiratory tract infection [22]. There is limited application of autoimmune profiling and human leukocyte antigen (HLA) profiling. However, HLA profiling might be helpful in managing suspected drug (allopurinol, carbamazepine, lamotrigine, phenobarbital, phenytoine, sulfamethoxazole, oxycam, and nevirapine) induced Stevens–Johnson syndrome and toxic epidermal necrolysis by confirming the HLA-B\*1502 and HLA-B\*5801 status in patients of Chinese (Han origin and Taiwan–Chinese) and people from other far east countries [23].



The scope of genetic- or genome-based diagnosis in acute or chronic inflammatory disorders is very limited. However, specific single-gene or multiple-gene panel next-generation genome sequencing (NGS) laboratory methods are used in making specific rare Mendelian (single-gene) diseases. Examples include cystic fibrosis, severe combined immune deficiency disease, and alpha-1-antitrypsin deficiency syndrome.

There is now an emerging approach to investigate patients with complex inflammatory diseases using whole-exome sequencing or whole-genome sequencing (WGS) next-generation technologies. A recent report [24] describes the outcome of WGS in systemic autoinflammatory disorders (SAIDs), a heterogeneous group of monogenic diseases sharing a primary dysfunction of the innate immune system. More than 50% of patients with SAID do not show any mutation at gene(s) tested because of lack of precise clinical classification criteria and/or incomplete gene screening. This study selected 10 genes (*MEFV*, *MVK*, *TNFRSF1A*, *NLRP3*, *NLRP12*, *NOD2*, *PSTPIP1*, *IL1RN*, *LPIN2*, and *PSMB8*) and undertook the molecular diagnosis and genotype interpretation of SAIDs in 50 patients.

The study concludes that the molecular diagnosis of SAIDs is possible using the multigene NGS technology. However, the study found genotype–phenotype correlation remarkably difficult, thus making this approach clinically not useful, supporting the commonly held view to apply the evidence-based and validated clinical criteria concurrently with the genetic analysis for the final diagnosis and classification of patients with SAIDs.

Previous attempts using the genome-wide association study (GWAS) to find the clinically applicable single-nucleotide polymorphisms (SNPs) and copy number variation across the genome have not led to any specific data in relation to complex chronic inflammatory diseases, for example, RA, inflammatory bowel disease (ulcerative colitis and Crohn's disease), and bronchial asthma.

## 22.5.2 Chronic inflammatory connective tissue diseases

Inflammatory connective tissue diseases predominantly involve joints, ligaments, synovial tissue, and skin. The term *arthritis* includes a large number of diseases in which the predominant feature is synovial inflammation. The three main subgroups of inflammatory arthritis are RA, spondyloarthritis, and crystal arthritis. The diagnosis of these conditions is helped by distinguishing the following clinical features:

- The pattern of joint involvement (symmetrical or asymmetrical, large or small)
- The presence of any nonarticular disease (psoriasis, iritis, IBS, nonspecific urethritis)
- A past and family history
- Periodicity of the arthritis (single acute, relapsing, chronic, and progressive)

Detailed discussion on individual types is outside the scope and remit of this chapter. However, relevant clinical information, pathogenesis, and pertinent molecular information are included here on predominantly articular (RA) and nonarticular (SLE) chronic inflammatory diseases. Both disorders share overlapping clinical features triggered with underlying autoimmune molecular mechanisms.

### 22.5.2.1 Rheumatoid arthritis

RA is an autoimmune disease associated with autoantibodies to the Fc portion of immunoglobulin G [rheumatoid factor (RF)] and to citrullinated cyclic peptide. Clinically it is characterized with persistent synovitis, causing chronic symmetrical polyarthritis with systemic inflammation. It is a heterogeneous condition following the polygenic/multifactorial inheritance.

#### 22.5.2.1.1 Genetic factors

The inheritance pattern in RA satisfies the criteria for polygenic/multifactorial genetic causation. There is a recurrent pattern of family history of an increased incidence in first-degree relatives. Twin studies consistently indicate a high concordance among monozygotic twins (up to 15%) and dizygotic twins (3.5%). Overall, genetic factors account for about 60% of disease susceptibility. There is a strong association between susceptibility to RA and certain HLA haplotypes, such as HLA-DR4, which occurs in 50%–75% of patients and correlates with a poor prognosis, as does possession of certain shared alleles of HLA-DRB1\*04. The possession of these shared epitope alleles in HLA-DRB1 (S2 and S3P) increases susceptibility to RA and may predispose to anticitrullinated peptide antibodies (ACPA) directed against citrullinated antigens. Citrullination is a process that modifies antigens, allowing them to fit into the shared epitope on HLA alleles. In a GWAS in ACPA-positive RA, an association

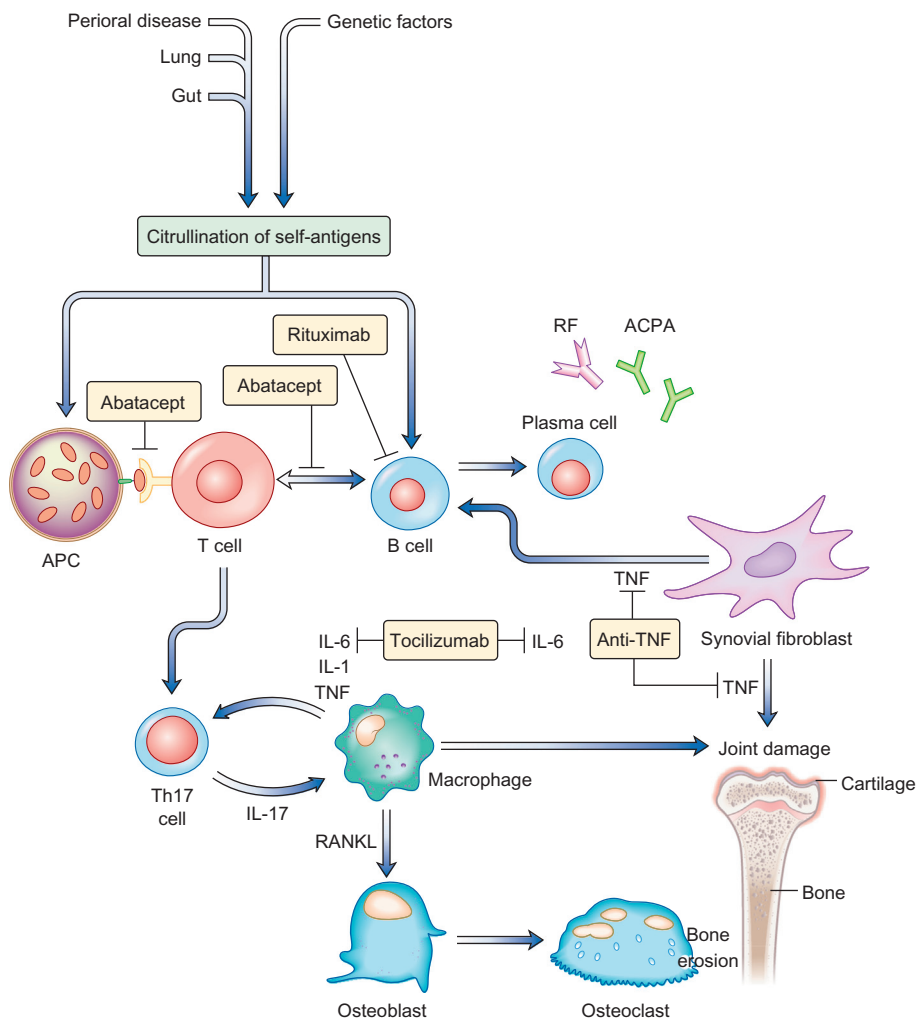
was found with loci near HLA-DRB1 and PTPN22 in people of European descent. These genes affect the presentation of autoantigens (*HLA-DRB1*), T-cell receptor signal transduction (*PTPN22*), and targets of ACPA (*PAD14*).

Eyre et al. [26] carried out an extensive metaanalysis of GWAS, using the SNP array. The group identified 14 new RA susceptibility loci, 5 of which were specifically associated with disease that was positive for ACPA, bringing the number of confirmed RA risk loci in individuals of European ancestry to 46.

### 22.5.2.1.2 Molecular pathology

The molecular pathology of RA has elements of autoimmune-led inflammatory process. The complex antigen–antibody interaction promotes overproduction of TNF- $\alpha$ , and then leads to synovitis and joint destruction. Interaction of macrophages and T and B lymphocytes drives this overproduction. TNF- $\alpha$  stimulates overproduction of IL-6, as well as other cytokines. The increased understanding of the immunopathogenesis of this disease has informed the development of targeted biological therapies (Fig. 22.4). Blockade of TNF- $\alpha$  and IL-6 has produced marked improvement in synovitis and systemic malaise, indicating the pivotal role of these cytokines in chronic synovitis.

The autoimmune aspects of RA involve RF and ACPA. Transient production of RF is an essential normal phenomenon for removing immune complexes. In RA, immune complexes show a much higher affinity to synovial tissue. They are of any immunoglobulin class (IgM, IgG, or IgA), but the most common tests employed clinically detect IgM RF. Around 70% of people with polyarticular RA have IgM RF in the serum. Positive titers can predate the onset of RA. The term “seronegative RA” is used when the standard tests for IgM RF are persistently negative. These patients tend to have a more limited pattern of synovitis.



**FIGURE 22.4** Pathogenesis of rheumatoid arthritis: environment–gene interactions promote citrullination of self-proteins, which can then be detected by T and B cells; this leads to a loss of tolerance and promotion of the inflammatory response, resulting in joint damage. Targeted therapy is also shown. ACPA, Anticitrullinated peptide antibody; APC, antigen-presenting cell; IL, interleukin; RANKL, receptor activator of nuclear factor kappa B ligand; RF, rheumatoid factor; Th, T helper; TNF, tumor necrosis factor. Source: Adapted with permission from Fig. 18.19, page 673, *Clinical Medicine* eds. Kumar and Clark, Churchill Livingstone Elsevier.

### 22.5.2.1.3 Clinical subtypes

The natural history and clinical manifestations may assist the clinician to put an RA patient in one of the following categories. However, it is a continuum without any firm distinguishing features:

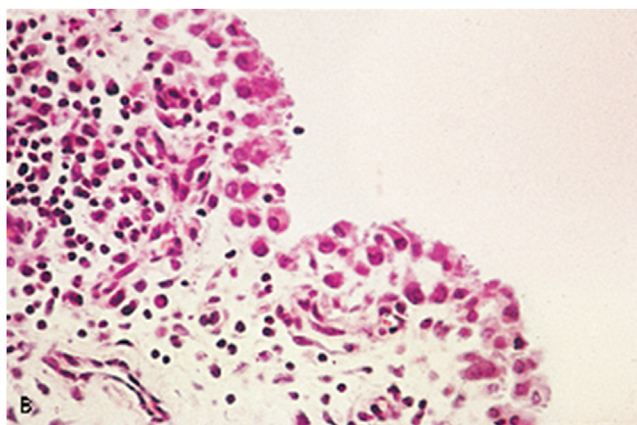
1. *Palindromic*—monoarticular that may progress to other types of RA.
2. *Transient*—self-limiting with permanent joint damage; usually seronegative for IgM RF and ACPA. Some of these patients may have had history of recent viral infection.
3. *Remitting*—the arthritis is active for several years but then remits, leaving minimal damage.
4. *Chronic, persistent*—most typical form, seropositive or seronegative for IgM RF. The disease follows a relapsing and remitting course over many years. Seropositive (plus ACPA) patients tend to develop greater joint damage and long-term disability. This type warrants earlier and more aggressive treatment with disease-modifying agents.
5. *Rapidly progressive*—the disease may progress remorselessly and leads rapidly to severe joint damage and disability; usually seropositive (plus ACPA) has a high incidence of systemic complications and is difficult to treat.

### 22.5.2.1.4 Articular features

The joint involvement in RA follows a unifying pattern in any of the above clinical subtypes. The synovium becomes greatly thickened, causing “boggy” swelling around joints and tendons, with proliferation of the synovium into folds and fronds, and infiltration by a variety of inflammatory cells. The normally sparse surface layer of lining cells becomes hyperplastic and thickened (Fig. 22.5). There is marked vascular proliferation. Increased permeability of blood vessels and the synovial lining layer leads to joint effusions that contain lymphocytes and dying polymorphs. The hyperplastic synovium spreads from the joint margins on to the cartilage surface. This “pannus” of inflamed synovium damages the underlying cartilage by blocking its normal route for nutrition and by the direct effects of cytokines on the chondrocytes. The cartilage becomes thin, and the underlying bone gets exposed. Local cytokine production and joint disuse combine to cause juxta-articular osteoporosis during active synovitis.

Fibroblasts from the proliferating synovium also grow along the course of blood vessels between the synovial margins and the epiphyseal bone cavity and damage the bone. This process is shown by MRI to occur in the first 3–6 months following onset of the arthritis before the diagnostic, ill-defined, juxta-articular bony “erosions” appear on X-ray.

In most patients the soft tissue involvement surrounding joints manifest with *subcutaneous nodules*. These are firm and intradermal, generally occurring over pressure points: typically, the elbows, the finger joints, and the Achilles tendon in patients with seropositive erosive disease. The olecranon and other bursae may be swollen (bursitis). *Tenosynovitis* of flexor tendons in the hand can cause stiffness and occasionally a trigger finger. Swelling of the extensor tendon sheath over the dorsum of the wrist is common. *Muscle wasting* around joints is



**FIGURE 22.5** Histological appearance of synovium in RA. Synovial appearances in established RA, showing marked hypertrophy of the tissues with infiltration by lymphocytes and plasma cells. RA, Rheumatoid arthritis. Source: Adapted with permission (Fig. 18.20, page 674) from Kumar and Clark's *Clinical Medicine*, Churchill Livingstone, Elsevier.

**TABLE 22.2** Nonarticular manifestations in rheumatoid arthritis (RA).

Organ/System
<ul style="list-style-type: none"> <li>Main clinical manifestations</li> </ul>
<b>Lungs</b> <ul style="list-style-type: none"> <li>Bronchiectasis (cough and daily sputum)</li> <li>Obliterative bronchiolitis (progressive breathlessness)</li> <li>Pleural thickening and effusion</li> <li>Interstitial lung disease: a combination of inflammation and basal lung fibrosis</li> <li>Peripheral, intrapulmonary nodules (Caplan syndrome)</li> <li>Infective lesions, for example, tuberculosis in patients on biological DMARDs</li> </ul>
<p><i>Vasculitis</i> caused by immune complex deposition in arterial walls is uncommon. Smoking is a risk factor. Findings include</p> <ul style="list-style-type: none"> <li>nail-fold infarcts due to cutaneous vasculitis</li> <li>widespread cutaneous vasculitis with necrosis of the skin (seen in people with very active, strongly seropositive disease)</li> </ul> <p><i>Cardiovascular</i>: Poorly controlled RA with hypertension and a persistently raised CRP, high cholesterol is a cardiovascular risk factor. Cardiovascular problems in RA may include</p> <ul style="list-style-type: none"> <li>pericarditis, which is rarely symptomatic</li> <li>endocarditis and myocardial disease, rarely symptomatic, (found at postmortem in approximately 20% of cases)</li> </ul>
<b>Nervous system</b> <ul style="list-style-type: none"> <li>Peripheral sensory neuropathies: mononeuritis multiplex or symmetrical, peripheral—due to vasculitis of the vasa nervorum</li> <li>Compression neuropathies: carpal or tarsal tunnel syndrome—due to synovitis</li> <li>Cord compression: due to atlantoaxial subluxation (see earlier)</li> <li>Raynaud syndrome (rare)</li> </ul>
<b>Eyes</b> <ul style="list-style-type: none"> <li>Sicca syndrome causes dry mouth and eyes</li> <li>Scleritis and episcleritis occur in severe, seropositive disease, resulting in painful red eye</li> <li>Scleromalacia perforans is a rare feature</li> </ul>
<b>Kidneys</b> <p>Proteinuria, nephrotic syndrome, and chronic kidney disease caused by renal amyloidosis. It is due to the deposition of highly stable SAP in the intercellular matrix of a variety of organs. SAP is a hepatic acute-phase reactant</p>
<b>Blood</b> <ul style="list-style-type: none"> <li><i>Felty syndrome</i>—splenomegaly and neutropenia in a patient with RA. Leg ulcers and sepsis are complications. HLA-DR4 is found in 95% of patients, compared with 50%–75% of people with RA alone. Lymph nodes may be palpable, usually proximal to affected joints. There may be peripheral lymphedema of the arm or leg</li> <li><i>Anemia</i>—almost universal, usually normochromic and normocytic. It may be iron deficient owing to gastrointestinal blood loss from non-steroidal anti-inflammatory drug (NSAID) ingestion, or rarely hemolytic (Coombs-positive). There may be a pancytopenia due to hypersplenism in Felty syndrome or as a complication of DMARD treatment. A high platelet count occurs with active disease</li> </ul>

CRP, C-reactive protein; DMARD, Disease-modifying antirheumatic drug; SAP, serum amyloid A protein.

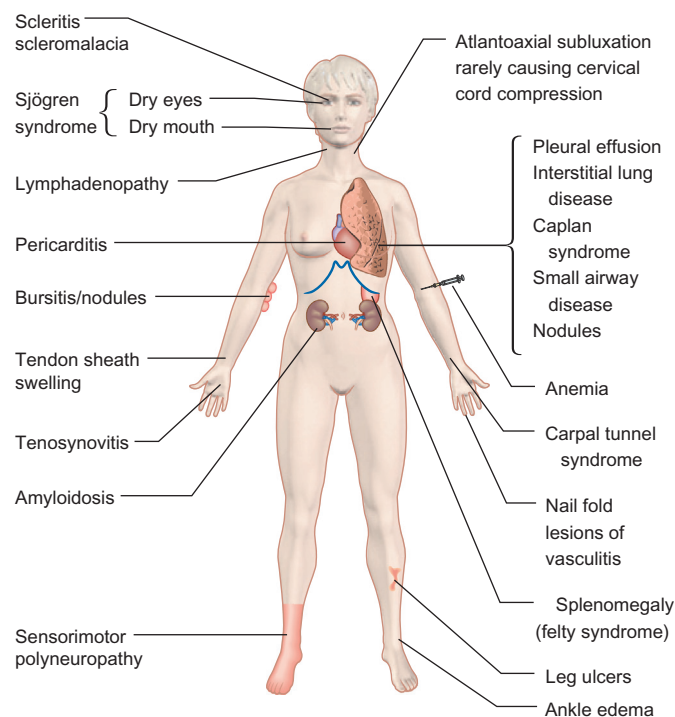
common complicated with corticosteroids. *Osteoporosis* is more common in poorly controlled RA and could be complicated with steroids (Table 22.2).

#### 22.5.2.1.5 Nonarticular manifestations

The clinical spectrum of RA includes several nonarticular features (Fig. 22.6). Some of these could manifest early, but most of the nonarticular manifestations are late. Clinical heterogeneity of nonarticular manifestations in RA makes this condition essentially a *multisystem chronic inflammatory* disease (Table 22.3).

#### 22.5.2.1.6 Treatment

The medical treatment of RA involves multiple pharmacotherapy to control the early articular damage. In most cases the early use of disease-modifying antirheumatic drugs within 3 months of onset of the arthritis is recommended to try to induce disease remission. Additional low-dose steroids may delay, and anti-TNF- $\alpha$  agents halt and occasionally reverse RA erosion formation. Erosions lead to a variety of deformities and contribute to long-term disability. The medical treatment needs to be combined with personalized physiotherapy and lifestyle changes.



**FIGURE 22.6** Nonarticular systemic manifestations in rheumatoid arthritis. Source: Adapted with permission (Fig. 18.24, page 677) from Kumar and Clark's *Clinical Medicine*, Churchill Livingstone, Elsevier.

**TABLE 22.3** Antinuclear antibodies in connective tissue autoimmune inflammation.

Antibody	Disease	Prevalence
Ds-DNA	SLE	70%
Antihistone	Drug-induced lupus	—
Anticentromeric	Limited scleroderma	70%
AntiRo (SS-A)	SLEPrimary Sjögren's syndrome	40% – 60% 60% – 90%
AntiLa (SS-B)	SLE Primary Sjögren's syndrome	15% 35% – 85%
AntiSm	SLE	10% – 25% (Caucasian) 30% – 50% (Black African)
AntiUI-RNP	SLE Overlap syndrome	30%
AntiJo-1 (antisynthetase)	Polymyositis Dermatomyositis	30%
Antitopoisomerase (Scl-70)	Diffuse cutaneous SSc	30%

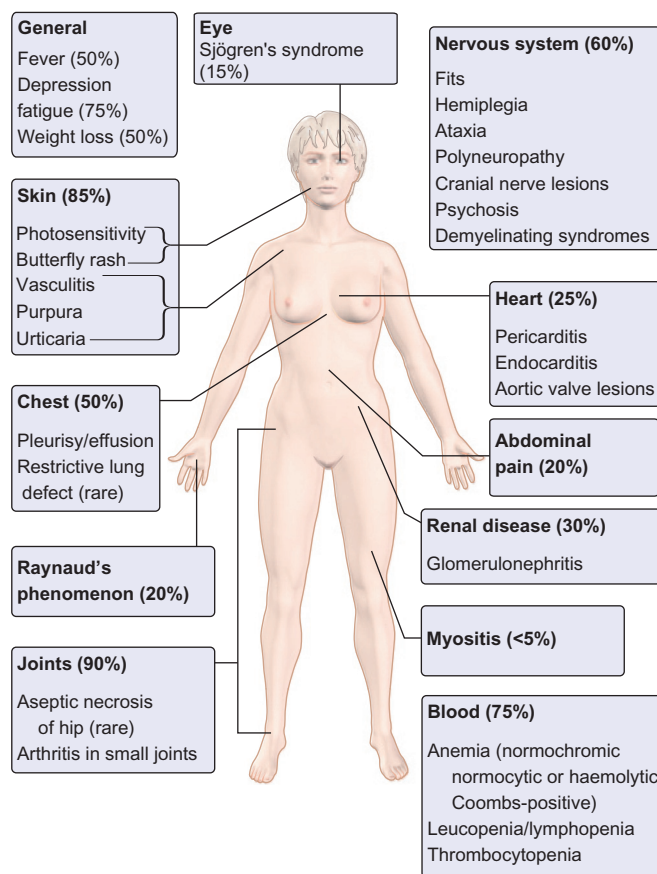
*ds-DNA*, Double-stranded DNA; *RNP*, ribonucleoprotein; *Ro*, *La*, first two letters of name of patients; *Sm*, Smith, patient's name; *SS-A*, *SS-B*, Sjögren's syndrome A and B; *SSc*, systemic scleroderma; *SLE*, systemic lupus erythematosus.

Adapted with permission from Box 18.36 page 694, *Clinical Medicine* by Kumar and Clark (eds), Churchill Livingstone Elsevier.

### 22.5.2.2 Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is a multisystem auto-immune chronic inflammatory disease with acute exacerbations. The age at onset and the natural history are extremely variable that form the full SLE clinical spectrum. In addition to range of articular and nonarticular inflammation, it is also characterised with progressive neurological, cardiovascular, and renal complications (Fig. 22.7).





**FIGURE 22.7** Clinical features of SLE. SLE, Systemic lupus erythematosus. Source: Adapted with permission from Fig. 18.35, page 693, Clinical Medicine eds. Kumar and Clark, Churchill Livingstone Elsevier.

### 22.5.3 Targeted molecular therapy for acute or chronic inflammatory diseases

Understanding the molecular biology of inflammatory processes has led to many new drugs capable of eliminating the senescent cells or suppressing the SASP [25]. This approach is one of the many similar attempts to find effective treatment for chronic inflammatory diseases. In animal studies, suppression of the Nlrp3 inflammasome-dependent proinflammatory cascade resulted in reduction of age-related degenerative changes across multiple organs [21]. Another prospective target is IL-6 that is considered important in local and/or systemic inflammation. New drugs and vaccines are under development, which are capable of neutralizing IL-6 and thus limiting proinflammatory cytokines causing widespread organ or tissue damage. Apart from aspirin, there are now a range of nonsteroidal antiinflammatory drugs available effective in treatment of patients with a range of chronic inflammatory diseases ([www.bnf.org](http://www.bnf.org)). In addition, lifestyle changes, modest exercise, and other simple physical measures might help in nonpharmacological suppression of the impact of in vivo inflammatory processes triggered by an injury, acute or chronic infection, autoimmune process, or age-dependent chronic inflammation.

## 22.6 Summary

The molecular basis of acute and chronic inflammation is complex with several underlying genetic and genomic factors. Main triggers of both acute and chronic inflammatory process include acute, recurrent, or chronic tissue damage resulting from trauma, sepsis (bacterial infection), persistent or chronic recurrent infection (bacterial, viral, fungal), immune-mediated *auto* or *isoantigen:antibody* reaction and neoplastic proliferation (localized or systemic). Age-associated (senescence) inflammation (inflammaging) is complex with complex molecular processes.

Inflammatory markers provide a useful index of both acute and chronic inflammation but are not usually specific with limited diagnostic accuracy. Nevertheless, these are important in clinical practice, particularly in acute critical care settings. Commonly used inflammatory markers include acute-phase reactants (LBP, CRP, procalcitonin); mediator activity (TNF, IL-1, IL-6, IL-10, IL-18) and cellular activity (TNF-RI, TNF-RII, IL-1R, sIL-6-R, mIL-6- $\alpha$ , ICAM-1, E-selectin, CD11b, elastase, HLA-DR class-II antigens, DNA). It is likely that the diagnostic and/or predictive power of these biomarkers could vastly improve when combined with genomic variants, such as SNPs and copy number variants. There is yet an unclear scope of WGS in managing acute or chronic inflammatory diseases. However, it is likely that novel antiinflammatory drugs could be discovered from the proteomic applications of complex inflammation genomic research.

### Acknowledgments and disclaimer

The author has used text, illustrations, data from several published resources including that exist on Google and other public domains. Limited copyright permission is secured. However, it is likely that some material could be subject to copyright approval for which the author or publisher are unaware. In any case the author acknowledges and expresses gratitude to all people, corporations, or any other organization or institute holding such rights for anticipatory permission to use such material on behalf of the Genomic Medicine Foundation (United Kingdom), a registered “not for profit” company in England and Wales.

### References

- [1] Harlan JM. Neutrophil-mediated vascular injury. *Acta Med Scand* 1987;715(Suppl):123–9.
- [2] Giannoudis PV. Current concepts of the inflammatory response after major trauma: an update. *Injury* 2003;34:397–404.
- [3] Giannoudis PV, Hildebrand F, Pape HC. Inflammatory serum markers in patients with multiple trauma—can they predict outcome? *J Bone Joint Surgery (Br)* 2004;86-B(3):313–23.
- [4] Smith RM, Giannoudis PV. Trauma and immune response. *R Soc Med* 1998;91:417–20.
- [5] Giannoudis PV, Smith RM, Banks RE, et al. Stimulation of inflammatory markers after blunt trauma. *Br J Surg* 1998;85:986–90.
- [6] Cipolle MD, Pasquale MD, Cerra FB. Secondary organ dysfunction: from clinical perspectives to molecular mediators. *Crit Care Clin* 1993;9:261–97.
- [7] Anderson BO, Harken AH. Multiple organ failure: inflammatory priming and activation sequences promote autogenous tissue injury. *J Trauma* 1990;30(Suppl):44–7.
- [8] Giannoudis PV, Smith RM, Ramsden CW, et al. Molecular mediators and trauma: effects of accidental trauma on the production of plasma elastase, IL-6, sICAM-1 and sE-selectin. *Injury* 1996;27:376–7.
- [9] Sablotzki A, Borgermann J, Bauling W, et al. Lipopolysaccharide-binding protein (LBP) and markers of acute phase response in patients with multiple organ dysfunction syndrome (MODS) following open heart surgery. *Thorac Cardiovasc Surg* 2001;49:273–8.
- [10] Seki E, Brenner DA, Karin M. A liver full of JNK: Signaling in regulation of cell function and disease pathogenesis, and clinical approaches. *Gastroenterology*, 2012;143(2):307–20.
- [11] Koyama Y, Brenner DA. Liver inflammation and fibrosis. *J Clin Invest* 2017;127(1):55–64. Available from: <https://doi.org/10.1172/JCI88881>.
- [12] Jiang J, Li J, Yunxia Z, Zhu H, Liu J, et al. The role of prostatitis in prostate cancer: meta-analysis. *PLoS One* 2013;8(12):e85179. Available from: <https://doi.org/10.1371/journal.pone.0085179>.
- [13] Schetter AJ, Heegaard NH, Harris CC. Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. *Carcinogenesis* 2010;31:37–49. Available from: <https://doi.org/10.1093/carcin/bgp272>. PubMed: 19955394.
- [14] Bollrath J, Greten FR. IKK/NF-kappaB and STAT3 pathways: central signalling hubs in inflammation-mediated tumour promotion and metastasis. *EMBO Rep* 2009;10:1314–19. Available from: <https://doi.org/10.1111/j.1749-6632.2009.tb06651.x>. PubMed: 19893576.
- [15] Franceschi C, Camisi J. Chronic inflammation (Inflammaging) and its potential contribution to age-associated diseases. *J Gerontol A Biol Sci Med Sci* 2014;69(S1):S4–9. Available from: <https://doi.org/10.1093/gerona/glu057>. For permissions, please e-mail. [journals.permissions@oup.com](mailto:journals.permissions@oup.com).
- [16] Brenner, C, Galluzzi L, Kepp O, Kroemer G, Decoding cell death signals in liver inflammation, *J Hepatol* 2013;59(3): 583–94.
- [17] Franceschi C, Bonafè M, Valensin S, et al. Inflamm-aging. An evolutionary perspective on immunosenescence. *Ann N Y Acad Sci*. 2000;908:244–54. Available from: <https://doi.org/10.1111/j.1749-6632.2000.tb06651.x>.
- [18] Campisi J, d’Adda di Fagagna F. Cellular senescence: when bad (things happen to good cells. *Nat Rev Mol Cell Biol* 2007;8:729–40. Available from: <https://doi.org/10.1038/nrm2233>.
- [19] Shaw AC, Joshi S, Greenwood H, Panda A, Lord JM. Aging of the innate immune system. *Curr Opin Immunol* 2010;22:507–13. Available from: <https://doi.org/10.1016/j.coi.2010.05.003>.
- [20] Gallenga CE, Parmeggiani F, Costagliola C, Sebastiani A, Gallenga PE. Inflammaging: should this term be suitable for age related macular degeneration too? *Inflamm Res* 2014;63:105–7. Available from: <https://doi.org/10.1007/s00011-013-0684-2>.
- [21] Youm YH, Grant RW, McCabe LR, et al. Canonical Nlrp3 inflammasome links systemic low-grade inflammation to functional decline in aging. *Cell Metab*. 2013;18:519–32. Available from: <https://doi.org/10.1016/j.cmet.2013.09.010>.

- [22] Charalampous T, Richardson H, Kay GL, et al. Rapid Diagnosis of Lower Respiratory Infection Using Nanopore-Based Clinical Metagenomics. Cold Spring Harbor Laboratory ([www.biorxiv.org](http://www.biorxiv.org)). <<https://doi.org/10.1101/387548>>.
- [23] Lonjou C, Borot N, Sekula P, Ledger N, Thomas L, Halevy S, et al. A European study of HLA-B in Stevens-Johnson syndrome and toxic epidermal necrolysis related to five high-risk drugs. *Pharmacogenet Genomics*. 2008;18(2):99–107. Available from: <https://doi.org/10.1097/FPC.0b013e3282f3ef9c>.
- [24] Rusmini M, Federici S, Caroli F, et al. Next-generation sequencing and its initial applications for molecular diagnosis of systemic auto-inflammatory diseases. *Ann Rheum Dis* 2016;75:1550–7.
- [25] Tchkonina T, Zhu Y, van Deursen J, Campisi J, Kirkland JL. Cellular senescence and the senescent secretory phenotype: therapeutic opportunities. *J Clin Invest* 2013;123:966–72. Available from: <https://doi.org/10.1172/JCI64098>.
- [26] Eyre S, et al. Investigation of susceptibility loci identified in the UK rheumatoid arthritis whole-genome scan in a further series of 217 UK affected sibling pairs. *Arthritis and Rheumatology* 2004;50(3):729–35.

# Molecular basis of susceptibility and protection from microbial infections

Rashid Merchant<sup>1</sup>, Pezad Doctor<sup>1</sup> and Ami Varaiya<sup>2</sup>

<sup>1</sup>Department of Pediatrics, Nanavati Super Speciality Hospital, Mumbai, India <sup>2</sup>Department of Microbiology, Nanavati Super Speciality Hospital, Mumbai, India

## 23.1 Introduction

As “omics” have contributed to the identification of genetic susceptibility traits in cancer research, these techniques could be ultimately extrapolated with success to the field of infectious diseases. Furthermore, we are recently beginning to fully grasp the significance of the microbiota and its interactions with the mammalian immune system in defining susceptibility to infection. Molecular testing is now well integrated into the clinical microbiology laboratory and for some pathogens has revolutionized the diagnosis and management of infectious diseases. One of the early success stories for the power of molecular detection was in the diagnosis of herpes simplex virus (HSV)-associated encephalitis. In the mid-1990s testing for HSV DNA in cerebral spinal fluid proved to be a sensitive and specific alternative to brain biopsy, the previous gold standard for establishing this diagnosis. This was shortly followed by US Food and Drug Administration approval of the first molecular microbiology test, which was used for the detection of *Chlamydia trachomatis* from genital specimens. These early tests relied on polymerase chain reaction (PCR), but other amplification methods, such as branched DNA and transcription-mediated amplification, also came into widespread use. Soon enough, it became a widespread clinical practice to quantify the amount of virus directly from clinical specimens. By measuring human immunodeficiency virus (HIV) type 1 (HIV-1) RNA in plasma, clinicians have been able to design personalized treatment plans and cater to specific patient needs like never before. The use of viral load testing has progressed and been applied to the management of hepatitis C virus (HCV) infections, as well as cytomegalovirus, Epstein–Barr virus (EBV), and Brennan-Krohn (BK) virus infections in bone marrow and solid organ transplant recipients.

The emergence of molecular testing in the clinic has been made possible by major advances in technologies, most notably the development of real-time methods that have enabled the simultaneous amplification and detection of target nucleic acids. These methods have reduced the turnaround time for testing to a few hours. There are now systems available, and others under development, that fully automate all steps of testing including nucleic acid extraction, amplification, and detection. The development of these systems has progressed along two paths. Large, high-throughput automated systems have been designed to increase efficiency in high-volume testing for HIV-1 and HCV viral load, as well as the detection of *C. trachomatis* and *Neisseria gonorrhoeae*. Meanwhile, smaller, simpler instruments allow the addition of the primary specimen directly into a cartridge or device, reporting results within a few hours. Moreover, these random-access instruments have on-board controls and do not require highly trained technologists [1].

## 23.2 Understanding host genetic variation of susceptibility to infectious disease

As clinicians look to further tap into the potential of molecular biology, host genetic variation could be the key. It is well established as a component of infectious disease pathogenesis and can give insight into potential treatments and long-term complications of infectious diseases. Candidate gene studies have highlighted many common, high-penetrance human genetic variants associated with infection and disease resistance. A deletion in chemokine receptor 5 (*CCR5*) confers resistance to certain strains of HIV, whereas a deletion in fucosyltransferase 2 (*FUT2*) confers resistance to symptomatic norovirus infection. Rare single-gene mutations, such as primary deficiencies of the Toll-like receptor (TLR) signaling system, specifically the *TLR3–Unc93b–TRIF–TRAF3* pathway, have been associated with increased susceptibility specifically to herpesvirus infection, particularly HSV encephalitis.

Over the last decade host genetic studies have rapidly progressed using single-nucleotide polymorphism (SNP) analysis, DNA sequencing, DNA microarrays, and cytogenetic methods. Recently, the technological development of high-throughput genotyping has provided a useful tool to examine the genetic basis of disease through genome-wide association studies (GWASs). This approach has considerably increased the number of known genes associated with major diseases and in some cases even opened the door to new treatments. For example, one of the first successful GWASs linked age-related macular degeneration to the *CFH* gene, which led to the discovery that inflammation drives the development of the disease. At the population level, understanding the genetic basis of susceptibility to infections will help us to pinpoint those who would benefit most from any future vaccines and may contribute to the development of novel treatment strategies [2].

### 23.2.1 Host genetic determinants of human immunodeficiency virus infection

Host genetic factors are important determinants of HIV-1 susceptibility, mother-to-child transmission (MTCT), and disease progression. HIV usually uses cluster of differentiation (CD)4 and a coreceptor to infect cells. The most common HIV coreceptors are the chemokine receptors *CCR5* and *CXCR4*. While most primary infections involve viruses that use *CCR5* as a coreceptor, *CXCR4* using virus is often identified in persons with more advanced disease and is associated with more rapid disease progression. To enter target cells, HIV interacts with the CD4 receptor via its gp120 protein, thereby stimulating a conformational change in gp120, which exposes a portion of transmembrane glycoprotein gp41, and allows access of the gp120 V-loop to either *CCR5* or *CXCR4*. Subsequently, a peptide in gp41 causes the fusion of the viral envelope and host cell membrane and allows the viral capsid to enter the target cell.

Identifying genetic variants that influence the response to HIV-1 can provide insights into approaches to predict disease progression, lead to development of new treatments, and provide new immunologic targets for vaccine development.

### 23.2.2 Coreceptor and their ligand variants in human immunodeficiency virus disease progression

Several SNPs within the cervical carcinomas (CC) chemokine receptor 5 (*CCR5*) coding and regulatory region appear to affect HIV-1 disease. *CCR5* is a critical coreceptor modulating perinatal transmission with a deletion of 32-bp from the coding region of the *CCR5* gene ( $\Delta 32$ ) providing almost complete protection against HIV-1 infection in individuals with the homozygous mutant genotype (*CCR5- $\Delta 32$ / $\Delta 32$* ). Individuals heterozygous for *CCR5-wt/ $\Delta 32$*  are less likely to be infected with HIV-1 and show a slower rate of disease progression. A polymorphism in the coding region of the CC chemokine receptor-2 (*CCR2*), a minor HIV-1 coreceptor, at position 180 (G to A polymorphism) has been associated with slower disease progression of HIV-1 infected adults.

### 23.2.3 Chemokine receptor genetic variants affecting HIV-1 mother-to-child transmission in absence of antiretrovirals

MTCT occurs predominantly with macrophage-tropic, nonsyncytium-inducing viruses that use *CCR5* as a coreceptor, and upregulation of *CCR5* expression in the placenta is associated with vertical transmission. Exposure to antiretrovirals modified the impact of these genetic variants on MTCT. Children with the



CCR5-59029-A allele, which has been associated with higher expression of CCR5, were less likely to be infected when exposed to nevirapine.

#### **23.2.4 Chemokine receptor genetic variants affecting HIV-1 mother-to-child transmission in absence of antiretrovirals**

There is a significant role for the CC chemokines CCL3 [macrophage inflammatory protein (MIP)-1 $\alpha$ ], CCL4 (MIP-1 $\beta$ ), and CCL5 (RANTES) in HIV-1 pathogenesis. CCL3 protein is encoded by two functional genes (CCL3/LD78 $\alpha$  and CCL3L1/LD78 $\beta$ ), occurring as two copies and as variable copy numbers, respectively, in different individuals. Infants born to HIV-infected mothers with lower levels of CCL3 production were at increased risk of infection.

#### **23.2.5 Innate immunity genetic associations with human immunodeficiency virus type 1 disease**

Mannose-binding lectin (MBL) protein, encoded by the *MBL2* gene, is an important determinant of the innate immune response during infection, and its deficiency has been identified as the cause of common opsonic defect of children. MBL deficiency has also been associated with increased HIV-1 vertical transmission, which emphasizes its important role in controlling HIV-1 infection in young infants and children.

#### **23.2.6 Human leukocyte antigen genotypes alter mother-to-child transmission and rate of disease progression**

The presence of B\*27 or B\*57 alleles was associated with slower disease progression which remained significant after adjustment for race, gender, age, and baseline HIV-1 log RNA, CD4<sup>+</sup> count and percent and weight for age z-score. Human leukocyte antigen (HLA) class I homozygosity has been associated with more rapid disease progression.

#### **23.2.7 Intracellular antiviral host factor affecting human immunodeficiency virus disease**

APOBEC3G (apolipoprotein B mRNA editing catalytic polypeptide 3G) is an endogenous inhibitor of HIV-1 replication strongly associated with more rapid decline of CD4<sup>+</sup> T cells and accelerated progression to AIDS.

#### **Conclusion**

Much research has established that host genetic factors are important determinants of HIV-1 MTCT and the rate of disease progression in children as well as adults. The judicious use of host genetics has the potential to improve the care and treatment of HIV-infected pregnant women and children.

From the treatment point of view, genetic variants in CYP genes and ABCB1 have been shown to alter the pharmacokinetics and response to specific antiretrovirals

and can be used to optimize treatment of children. HLA-B5701 allele is associated with the hypersensitivity reaction associated with abacavir. Hence, screening for this allele has become standard of care for persons prior to initiating antiretroviral therapy containing abacavir. Host genetics have provided new insights into innate and adaptive immunologic mechanisms used to control HIV-1 infection. It is likely that these approaches will lead to novel new strategies for vaccine development [3–6].

### **23.3 Clinical relevance of human leukocyte antigen gene variants in HBV infection**

Numerous genetic variants may influence the natural progression of hepatitis B virus (HBV) infection. The HLA system, the major histocompatibility complex (MHC) in humans, is an important host factor that is shown to correlate with the clinical course of HBV infection. GWASs have shown that SNPs near certain HLA gene loci are strongly associated with not only persistent HBV infection but also spontaneous HBV clearance and seroconversion, disease progression, and the evolution of liver cirrhosis (LC) and HBV-related hepatocellular carcinoma (HCC) in chronic hepatitis B (CHB).

### 23.4 Human leukocyte antigen gene variants and susceptibility and persistence of HBV infection

A comparative review showed that HLA-DRB111/12 alleles and DQB10301 are associated with HBV persistence globally. An association analysis performed across various ethnicity, including Japanese, Korean, Hong Kong, and Thai subjects, revealed Asian-specific associations of HLA-DPA1 and HLA-DPB1 alleles/haplotypes with HBV infection and disease progression, identifying a new risk allele HLA-DPB109:01 and a new protective allele DPB102:01 in chronic HBV infection. Persistence of HBV was shown to be associated with class II allelic homozygosity. Interestingly, three SNPs belonging to the HLA-DQ region (rs2856718, rs7453920, and rs9275572), shown to display increased susceptibility to chronic HBV infection, were found in Saudi Arabian patients.

#### 23.4.1 Human leukocyte antigen gene variants and spontaneous HBsAg clearance

Spontaneous HBsAg clearance seldom occurs in a small proportion of patients with chronic HBV infection. The mechanisms of spontaneous HBV clearance are influenced by the interactions between HBV and the host innate and adaptive immune responses, which are affected by specific HLA gene polymorphisms that change peptide epitope binding. HLA-DR13 is consistently associated with HBV clearance globally. A meta-analysis pointed out that patients possessing at least one A allele of HLA-DPB1 rs9277535 and HLA-DPA1 rs3077 variants have increased susceptibility to spontaneous HBV clearance compared with those with G alleles.

#### 23.4.2 Human leukocyte antigen gene variants and early HBeAg seroconversion

It is well known that HBeAg seroconversion depends on patient age at infection and the host immune responses. It was shown that the functional stage of dendritic cells (DCs) plays an important role in HBeAg seroconversion [7]. DCs are the most effective antigen presenting cells and play a pivotal role in antiviral response induction. A long-term cohort study showed that HLA class I antigen B61 and class II antigen DQB10503 are associated with early HBeAg seroconversion in CHB children in Taiwan.

#### 23.4.3 Human leukocyte antigen gene variants and risk of developing liver cirrhosis and HBV-related hepatocellular carcinoma

HLA gene variations are also strongly associated with disease progression and the development of LC and HBV-related HCC. Of the SNPs reported in HBV-related HCC GWASs, rs9267673 near C2, rs2647073 and rs3997872 near HLA-DRB1, and rs9275319 near HLA-DQ were noted to be significantly associated with the risk for HBV-related LC, suggesting that human genetic variants associated with HBV-related hepatocarcinogenesis may already play an important role in the progression from CHB to LC.

#### 23.4.4 Human leukocyte antigen gene variant and response to hepatitis B virus vaccine

A meta-analysis, including a total of 2308 subjects (1215 responders, 873 nonresponders, and 220 control populations) and assessing the effect of HLA on immunological response to hepatitis B vaccines in healthy individuals, showed that, for DRB1 alleles, the three HLA variants DRB101, DRB11301, and DRB115 are associated with significantly increased antibody response to hepatitis B vaccines, with pooled ORs of 2.73, 5.94, and 2.29, respectively.

#### 23.4.5 Human leukocyte antigen gene variants and efficacy of interferon alfa and NAs treatment

Interferon (IFN)- $\alpha$  is the first-line therapy for CHB patients but a complete response was established only in a minority of patients as HLA gene variants were shown to affect the response to IFN- $\alpha$  treatment in CHB patients. The “G–C” haplotype of the five SNPs, including rs9277535 (HLA-DPB1), rs9276370 (HLA-DQA2), rs7756516 and rs7453920 (HLA-DQB2), and rs9366816 near HLA-DPA3, were shown to be associated with sustained therapeutic response to IFN- $\alpha$  treatment in male Han Taiwanese subjects. The HLA-DQ locus rs9275572 predicts viral and biochemical responses to lamivudine (LAM) therapy in Han Chinese subjects. Hosaka et al. established an association of HLA-DP polymorphisms with  $\geq 2A$  alleles at rs3077 and rs9277535 and decreased HBsAg levels and seroclearance among HBeAg-positive Japanese CHB patients treated with LAM. While evaluating the impact of

HLA gene variants on HBV infection, SNP–SNP interactions between HLA and other host genes such as granulysin SNPs and polymorphisms in TLR–IFN pathway genes and HBV mutations should be considered.

### Conclusion

Advancement in our understanding of the molecular biology and replication cycle of HBV have provided unprecedented insight into the mechanisms of action

and treatment response of currently available drugs against HBV as well as potential future targets for therapeutic development [8,9].

### 23.4.6 Host genetic determinants in hepatitis C virus infection

HCV is an RNA flavivirus currently infecting approx. 170 million people worldwide [1]. Acute HCV infection is asymptomatic in much of the patients but persists in about 70% of them. These patients with persistent liver inflammation are at risk of disease progression to liver fibrosis, LC, and HCC with potentially fatal outcomes.

## 23.5 Immunogenetics and microbial infection

### 23.5.1 Genes involved in innate immunity

SNPs in the promoter region of the IFN regulatory factor-1 (IRF-1) were associated with protection from viral persistence. NK cells are lymphoid cells with the ability to exert antiviral functions through secretion of antiviral cytokines or lysis of infected cells. Importantly, killer cell immunoglobulin-like receptors expressed by NK cells interact with certain HLA class I molecules expressed by target cells.

### 23.5.2 Genes involved in adaptive immunity

HLA class I: In a large study by Thio et al. in Caucasians and black Americans, HLA-A1101 and HLA-B57 were found significantly more often in 231 individuals with well-documented HCV clearance compared to 444 matched chronically infected patients. HLA-B27 also plays a prominent role since the mechanisms of protection can be tracked down to the generation of HCV-specific CD8 + T-cell responses against a single viral epitope.

HLA class II: HLA-DQB1\*0301 was associated with clearance of HCV infection in studies conducted in populations throughout the world.

### 23.5.3 Genes involved in T-cell regulation and function

The function of CD8 + T cells depends on the maturation stage, which can be assessed by the combination of several differentiation markers linked to T-cell functions. Naïve T cells express a large isoform of the protein tyrosine phosphatase CD45, termed CD45RA. Upon activation, expression of this isoform is downregulated in T cells, and a short isoform, CD45RO, is expressed. The CD45 gene polymorphism C77G was more commonly seen in patients with HCV infection compared to the overall population. Genes responsible for T-cell regulation and function have been reported to affect the outcome of HCV infection. Specifically, polymorphisms involved in the suppression of T cell responses by interleukin (IL)-10 may affect the natural history of HCV infection.

### Conclusion

A better understanding of the role of the host genetic background in patients with HCV infection is pivotal for the development of new prophylactic and immunomodulatory antiviral strategies. With advances in the molecular understanding of crucial components of the viral life cycle, new direct-acting antiviral agents (DAAs) have been developed at a remarkable pace. Combining DAAs targeting different stages in the viral life cycle has proven

highly effective and enabled the development of IFN-free and largely ribavirin-free regimens, greatly improving the tolerability of therapy. With well-tolerated oral regimens, cure rates now exceed 90% for most patient populations. Inhibition of HCV replication has focused on three major viral targets: the NS3/4A protease (simeprevir), the NS5B RNA-dependent RNA polymerase (sofosbuvir), and the NS5A protein (ledipasvir, daclatasvir) [10–12].

## 23.6 Host genetic susceptibility to human papillomavirus infection and development of cervical cancer

Cervical cancer is the second most common cancer in women worldwide with an annual 493,243 cases and 273,505 deaths. The main cause of cervical cancer is infection by human papillomavirus (HPV). During their lifetime, many women become infected with HPV, but interestingly, only a small fraction of them develop cervical cancer and the rest regress to a normal healthy state. This suggests the role of additional risk factors playing an important role in the outcome of the infection. These risk factors include host and viral genetic factors along with environmental and lifestyle factors.

Host genetic risk factors to HPV infection and cervical cancer: The genetic link to cervical cancer development is strongly supported by epidemiological studies. A hereditary component of cervical tumors was detected in comparisons of twins and in a mother–daughter family study:

1. HLA genes: *HLA* A2, A\*01, A\*24, A\*1104, B7, B15, B63, and Cw\*0202 loci have been reported to be associated with increased risk of developing cervical cancer. On the other hand, DQw1, DQw3 alleles have been associated with increased risk of developing the disease.
2. TP53: *Tp53* gene localized on chromosome 17q 13.1 encodes protein 53 (p53) and is a tumor suppressor gene that regulates the cell cycle. HPV 16 and 18 encodes two major oncoproteins, E6 and E7. These two proteins interfere with the cellular tumor suppressor proteins, a seven times higher risk of developing HPV-associated cervical carcinogenesis in individuals with Arg homozygosity than with heterozygous genotype of TP53 polymorphism in codon 72. *Tp53* codon 72 Arg allele was also found to have decreased association to cervical abnormalities and HPV infection. On the other hand, the *Tp53* Pro/Pro genotype was reported to have a higher risk in developing cervical cancer.
3. Chemokine receptor-2 (CCR2): MCP-1 is a ligand for chemokine receptor CCR2 produced largely by tumor cells. MCP-1 is responsible for recruiting macrophages to tumors in bladder, cervix, ovary, lung, and breast. When the epithelial cells are infected by HPV, it reduces the MCP-1 expression from low-grade squamous intraepithelial lesions to high-grade squamous intraepithelial lesions (HSIL) and the levels of MCP-1 expression increases again from HSIL to invasive cervical cancer. A SNP of G to A at position 190 of CCR2 gene changes amino acid valine (GTC) to isoleucine (ATC) at codon 64 (CCR2-V64I). The A allele was also found to be a risk allele for developing HSIL and cervical cancer compared to healthy controls.
4. Fas and Fas ligand (FasL): Apoptosis is dependent greatly on signals from cell surface death receptor Fas/CD95. Reduced expression of FasL inhibits the apoptotic activity of the Fas–FasL pathway. An association of Fas-670A allele and A/A genotype with higher risk of developing cervical cancer was reported.

### Conclusion

Therapies based on the insertion of cytokine or other immunostimulatory genes into the genome of tumor cells followed by vaccination with the resulting, genetically altered, cytokine-producing vaccines represent a new potential strategy for the treatment of cancer patients. HPV 16 is the etiological agent of more than 60% human CC. Presently, two prophylactic vaccines against HPV 16 are available (GlaxoSmithKline “Cervarix” and Merck “Gardasil”). These vaccines can almost completely protect the immunized individuals against both, persistent HPV 16 infection and HPV 16–related cervical cytological pathology. However, no clinically utilizable therapeutic vaccines against CC are available.

During the last decade, animal models have substantially contributed to the development of the

therapeutic vaccines against HPV 16–associated tumors. It has been demonstrated that the HPV 16 E6/E7 oncoproteins can serve as tumor rejection antigens and that the HPV 16–associated tumor cells can be genetically modified with DNA encoding immunostimulatory cytokines [IL-2, IL-12, granulocyte-macrophage colony-stimulating factor (GM-CSF)] or other immunostimulatory molecules, used for vaccination, and inhibit tumor growth. To improve the HPV 16 antigen presentation in tumor-bearing individuals, DC-based vaccines loaded with HPV 16 E6/E7 DNA or hybrids of the dendritic and tumor cells have been successfully employed. Unfortunately, when these approaches used in animal models were applied into clinical trials, the results were less optimistic [13,14].

## 23.7 Host genetics of Epstein–Barr infection

EBV is a human gamma herpesvirus, infecting over 95% of adults by the age of 30 years. Childhood infections are usually clinically silent or difficult to distinguish from other mild viral infections. Infectious mononucleosis (IM) is usually a benign, self-limiting lymphoproliferative disease most common in the Western world, which occurs in between 25% and 70% of young adults following primary EBV infection. Symptoms of IM appear around a month after primary infection and range from benign (lymphadenopathy, sore throat, fever, and fatigue) to severe (fulminant hepatitis, liver necrosis, and/or hemophagocytic lymphohistiocytosis-HLH). The appearance and increase of EBV-specific T lymphocytes leads to a gradual decline in symptoms. Rarely, severe or fatal IM develops if no successful EBV-cytotoxic T lymphocyte response is mounted, which may be sporadic or linked to genetic disorders such as X-linked lymphoproliferative disease (XLP). Certain cancers and other chronic conditions are epidemiologically linked to IM, with independent epidemiological links between HL and MS following a clinical diagnosis of IM.

There has been greater success in identifying rare mutations associated with severe responses to primary EBV infection of B cells, of which XLP is perhaps the best known. XLP is a genetic disorder with a prevalence of approximately 1 in 1 million men. People with XLP are clinically healthy in 90% of cases until primary infection with EBV, which results in fulminant IM. Mutations in two human genes have so far been implicated in XLP: *SH2D1A* and *XIAP*. Some affected individuals do not carry these mutations, suggesting further genetic factors are involved in this disease.

Exome and genome sequencing have further identified four disorders involving genetic susceptibility to EBV infection: XMEN, *ITK* deficiency, *CORO1A* deficiency, and *PRKCD* deficiency.

To date, there are more than 30 host genes that have been associated with EBV infection, immunity, and disease. Those that are of potentially the greatest interest are the genes that have been associated with more than one EBV-related pathology or aspect of EBV immunity. Here we distill such genes and variants with disease overlap.

### 1. Fc fragment of IgG, low affinity IIa, receptor associated with Hodgkin's and non-Hodgkin's lymphoma

*FCGR2A* is present on the cell surfaces of macrophages, neutrophils, and NK cells, with roles in phagocytosis and modulation of the immune response. The studies linking this gene to EBV-positive B-cell lymphoma have both been relatively small and have examined a disparate collection of classical HL and non-HL, non-BL, making it difficult to draw conclusions on the role of *FCGR2A* in EBV susceptibility and disease. SNP rs1801274 is a functional polymorphism affecting binding of this IgG receptor, encoding a histidine to arginine substitution. The histidine-encoding allele is also associated with Kawasaki disease, an autoimmune disorder of unknown etiology in which EBV may play a role.

HLA system and association with infectious mononucleosis, multiple sclerosis, EBV antibodies, Hodgkin's lymphoma, posttransplant lymphoproliferative disorder, and nasopharyngeal carcinoma

Large and small studies of EBV antibody responses have identified association with polymorphisms with the HLA system (including classes I and II polymorphisms), which are so far independent from the polymorphisms identified as driving other EBV-related disease.

### 2. Perforin and association with chronic active EBV and hemophagocytic lymphohistiocytosis

*PRF1* permits granzymes A and B to reach the cytoplasm of cells targeted for destruction. *PRF1* mutations are found in 30% of HLH cases and may be essential for protection from herpes virus–driven lymphoproliferation. For example, Kaposi's sarcoma–associated herpesvirus-positive HLH occurred in two siblings with *PRF1* mutations. The association with GZMB polymorphisms and HLH suggests an essential role for targeted cell death in control of EBV infection.

### 3. Tumor necrosis factor alpha (TNF- $\alpha$ ) and association with posttransplant lymphoproliferative disorder and gastric carcinoma

Small studies of two different conditions (PTLD, GC) have linked host SNPs within *TNF* (a major inflammatory cytokine) to variable susceptibility to these EBV-related disorders. In vitro, EBV immediate-early lytic gene BZLF1 downregulates the TNF- $\alpha$  receptor, preventing TNF- $\alpha$ -induced cell death and signaling. However, no GWAS of EBV-related conditions have reported an association with *TNF* polymorphisms.



### Conclusion

To further our understanding of the mechanisms by which EBV causes pathology and predict which individuals are at greatest risk of EBV-mediated disease, an increased focus on host genetics of EBV infection is required. The methodology of these studies must shift from candidate genes interrogated for associations with a disparate range of disorders in small populations to

genome-wide approaches. Only large, well-powered studies can reliably identify common genetic polymorphisms contributing to the risk of EBV-related disease, with the hope of improving screening and treatment of these conditions. As the cost of such studies continues to fall, the next 50 years of EBV research may see more progress in this area [15–17].

## 23.8 Dengue viral infection

Dengue infections (DI) are a serious cause of morbidity and mortality in tropical and subtropical areas of the world including Southeast and South Asia, Central and South America, and the Caribbean. The disease is caused by dengue virus (DV), a *flavivirus* transmitted to humans mainly by infected *Aedes aegypti* mosquitoes. Infection by any of the four serotypes of DV, DV-1, -2, -3, and -4, may result in a wide clinical spectrum, ranging from asymptomatic to dengue fever (DF), dengue hemorrhagic fever (DHF), and dengue shock syndrome, the life-threatening complications being characterized mainly by plasma leakage.

The postulated factors exerting an influence on different manifestations can be divided into viral and host factors.

*Viral factors:* Higher blood viral load was found in DHF patients in comparison to DF patients. DV-2 had a larger pleural effusion index than the other virus serotypes in Thailand and clearly related to the severe clinical forms in a study in Vietnam.

*Host factors:* Plasma leakage often develops at day 4–6 of fever when the viremia has already declined, revealing the role of host immunopathological mechanisms in disease severity and progression. Both dengue virus–infected monocytes with antibody-dependent enhancement phenomenon, and activated specific T lymphocytes are responsible for the rapidly increased levels of cytokines in DHF. These cytokines, especially TNF- $\alpha$ , IFN-gamma (IFN- $\gamma$ ), and chemical mediators, including IL-1 and IL-6 from mast cell, play a key role in inducing important clinical manifestations of DHF, that is, plasma leakage and shock.

### 23.8.1 Dengue and major histocompatibility complex antigens

Many HLA class I alleles have been shown to be associated with severe dengue in secondary DI, suggesting the importance of the existing primed memory HLA class I–restricted cross-reactive T cell. Meanwhile, HLA class II, especially DRB1 alleles, more likely exerted a protective effect on DI and disease severity. A better understanding of this protection mechanism may lead to novel preventive and immunotherapeutic approaches, including vaccines. Located in the central or class III subregion of MHC region, TNF, an important vasoactive immunomodulator produced by activated monocytes, is known to be upregulated in DHF infections. Polymorphism in the promoter region of the TNF- $\alpha$  gene, –308A allele was found to be a risk for the development of DHF in South American patients. This SNP was reported to be important in diabetes mellitus, asthma, and allergic rhinitis most being associated with DHF.

### 23.8.2 Cytokine polymorphism and dengue

The production of cytokines is modulated by genetic polymorphisms that are associated with susceptibility to disease. IL-10 (–1082/–819/–592) ACC/ATA haplotype was significantly associated with DHF in a Cuban study (OR = 2.54,  $P_c$  = 0.03). Since IL-10 was involved with TNF- $\alpha$  in the thrombocytopenia and hemorrhagic manifestation in DI, the combination of TNF- $\alpha$  and IL-10 polymorphism is of interest. TNF- $\alpha$ -308 AA or AG and IL-10-1082 AA genotypes (high TNF/low IL-10 phenotype) were more frequent in DHF patients as compared with controls (OR = 19.47) or DF (OR = 17.4). Hence, many combinations of TNF- $\alpha$ , IFN- $\gamma$ , TGF $\beta$ 1 polymorphisms and IL-10 haplotype were found to be associated with DHF as compared with controls by Perez in 2010.

The most effective way to control dengue diseases in the future will include the use of a safe and effective vaccine. Vaccine developers have focused on the generation of a tetravalent vaccine aimed at providing

long-term protection against all virus serotypes. The first dengue vaccine, Dengvaxia (CYD-TDV) by Sanofi Pasteur, was first registered in Mexico in December 2015. CYD-TDV is a live recombinant tetravalent dengue vaccine that has been evaluated as a three-dose series on a 0/6/12 month schedule in Phase III clinical studies. It has been registered for use in individuals 9–45 years of age residing in endemic areas. There are approximately five additional vaccine candidates under evaluation in clinical trials, including other live-attenuated vaccines, as well as various subunit, DNA, and purified inactivated vaccine candidates [18,19].

## 23.9 Genetic susceptibility of humans to hantavirus infection

The clinical course of human hantaviral infections varies greatly according to different hantaviruses, ranging from no disease to mild course and low case-fatality rate (0.1% in *Puumala* virus—PUUV—infection) to severe course up to 40%–50% in *Sin Nombre*—SNV—and *Andes*—ANDV—virus infections. In addition, large variation in clinical severity occurs among patients for a given hantavirus species.

### 23.9.1 Immunity-related gene polymorphisms and severity of hantavirus infections

Gene candidate approaches have identified associations between human genotypes and the clinical severity of hantavirus infections. For hemorrhagic fever with renal syndrome and HCPS (hantavirus cardiopulmonary syndrome), risk *HLA* haplotypes have been identified according to the following clinical and laboratory parameters of disease severity: treatment time at hospital (overall severity), weight change during hospital care (amount of fluid retention during oliguric phase), need of dialysis, lowest systolic blood pressure, presence of shock, increase of plasma creatinine and urea (severity of acute kidney injury), thrombocytopenia, and leukocytosis.

In Finland, patients with *HLA* alleles HLA-B\*08 and DRB1\*0301 were more likely to have the most severe form of the PUUV infection with lower blood pressures, higher creatinine, and more virus excretion into the urine and into the blood. On the other hand, individuals with HLA-B\*27 have a benign clinical course.

In the United States the HLA-B\*3501 and HLA-DRB1\*1402 alleles are associated with increased risk of severe SNV-induced HCPS. HLA-B\*35-restricted memory T-cell responses were related to mild disease outcome in HCPS due to ANDV. Hence, different hantaviruses seem to be processed differently through the same *HLA* molecules resulting in mild or severe course of the disease.

### 23.9.2 Immune-related gene expression variability and severity of hantavirus infection

Several associations between serum levels of cytokines TNF IL-6, IL-2, IL-8, IL-10, IFN- $\gamma$ , or the intensity of platelet  $\beta 3$  integrin and disease severity have been identified for PUUV, Hantaan virus (HTNV), and Dobrava hanta virus (DOBV) infections. Genetic determinisms modulating the mRNA expression levels of the genes encoding these molecules could represent important risk factors of hantavirus disease severity [20].

### 23.9.3 Genetic susceptibility to severe influenza

Mouse studies clearly demonstrate that host genetics plays an important role in susceptibility to a range of human and avian influenza viruses. The *Mx* genes encoding IFN inducible proteins are the best studied, but their relevance to susceptibility in humans is unknown. One genealogy study shows moderate evidence of a heritable component to the risk of influenza-associated death, and while the marked familial aggregation of H5N1 cases is suggestive of host genetic factors, this remains unproven. The on-going family clustering of highly pathogenic avian influenza A/H5N1 cases, as demonstrated by the deaths in 2011 of a mother and son in Cambodia, and of two siblings and their mother in Indonesia, has proposed that host genetics play a role in susceptibility to H5N1 influenza. Even though H5N1 is uniquely virulent influenza virus, patterns of disease in other influenza epidemics also suggest a possible role for host genetics in susceptibility to severe influenza. Although viral genetic determinants of influenza severity have been intensively studied, host determinants are much less studied.

### 23.9.4 Influenza-associated encephalopathy

Acute encephalitis is a rare but a well-recognized complication of influenza infection, which occurs mostly in children aged under 5 years and is reported more commonly in East Asia than elsewhere. There is little data to assess if there is genetic susceptibility to influenza-associated encephalopathy (IAE) other than a report of a mother and daughter with H1N1/09 IAE, two siblings with H5N1 IAE, and an analysis of three IAE cases, which described a missense mutation in the TLR3 gene in one case. Acute necrotizing encephalopathy (ANE) is a distinct clinical syndrome that is characterized by multiple necrotic brain lesions and is associated with viral infections including influenza. A subset of patients with recurrent or familial ANE (ANE1) have a missense mutation in the ran-binding protein 2 (RANBP2) gene on chromosome 2 (q12.3).

#### Conclusion

Hitherto, the question “Is susceptibility to severe influenza in humans heritable?” remains unanswered. A better understanding of the biological predispositions

and pathways leading to severe influenza may lead to improved therapeutic options [21,22].

## 23.10 Host factors and genetic susceptibility to intracellular bacteria

### 23.10.1 *Mycobacterium tuberculosis*

Tuberculosis remains a major public health problem. In 2015, the incidence of new TB cases were 10.4 million and 1.8 million died from the disease (including 0.4 million among people with HIV). Over 95% of TB deaths occur in low- and middle-income countries. Although one-third of the world's population is exposed to TB, not all the individuals will become infected. Among those infected, only about 5% will develop clinical disease in their lifetime, suggesting the role of underlying genetic factors in TB. Increased prevalence of extrapulmonary tuberculosis among non-Caucasian populations and increased rates of TB disease among monozygote twins (60%) versus dizygotes twins (35%) support a major role for human genetic factors in the development of TB. Recent documentation of high variability rates of tuberculin skin test (TST) responsiveness and quantitative IFN- $\gamma$  release assay reactivity following exposure to active TB cases in healthy children supports the role of host factors.

Genetic variants correlated with increased susceptibility to TB include *IL10* promoter haplotypes and genome-wide linkage analysis (GWAS). Increased rates of an *IL10* promoter haplotype resulting in low levels of circulating IL-10 were found among TST-positive subjects compared with TST-negative ones, whereas GWAS correlated specific chromosomal regions (2q21–2q24 and 5p13–5q22) with persistent TST negativity.

While severe primary TB commonly occurs in children under 2 years of age, only a minority of infected children will develop severe forms of the disease, thus supporting the role of single inborn errors of immunity. Indeed, monogenic Mendelian defects in the mononuclear phagocyte/T helper cell type 1 (Th1) pathway were documented in up to 40% of children with disseminated primary TB. Among these, primary immunodeficiencies resulting from mutations in several genes, such as complete IFNGR1, IFNGR2, IL12B, IL12BR1, STAT1, IRF8, and ISG15L, were documented in children with disseminated TB disease. Polygenic somatic and germinal mutations, such as HLA polymorphisms (HLA-DR2 and HLA-DQB1), as well as mutations in Toll-like 1 and 2 receptors (TLR1, TLR2) and genes coding for the vitamin D receptor may be responsible for reactivation of pulmonary TB in adults, although no convincing evidence has been provided so far.

Mendelian susceptibility to mycobacterial disease (MSMD) is a condition characterized by predisposition to clinical disease caused by weakly virulent mycobacteria, such as bacille Calmette-Guerin (BCG) vaccines and environmental mycobacteria, in otherwise healthy individuals with no overt abnormalities in routine immunological tests. MSMD designation does not summarize all the clinical features, as patients are also prone to salmonellosis, candidiasis, and tuberculosis, and more rarely to infections with other intramacrophagic bacteria, fungi, or parasites, and even, a few viruses. Since 1996, several MSMD-causing genes, including autosomal (*IFNGR1*, *IFNGR2*, *STAT1*, *IL12B*, *IL12RB1*, *ISG15*, and *IRF8*) and two X-linked (*NEMO*, *CYBB*) genes have been discovered. These nine gene products are physiologically related, as all are involved in IFN- $\gamma$ -dependent immunity. These disorders impair either the production of (*IL12B*, *IL12RB1*, *IRF8*, *ISG15*, *NEMO*) or the response to (*IFNGR1*, *IFNGR2*, *STAT1*, *IRF8*, *CYBB*) IFN- $\gamma$ .

In contrast, various studies proposed the role of specific macrophage protein 1 (NRAMP1) gene polymorphisms in pulmonary TB, with a variable effect across populations (African, Asian populations vs European populations), epidemiological settings, clinical phenotypes, and age at onset of TB. A recent meta-analysis correlated NRAMP1 polymorphisms with pulmonary TB in African and Asian populations but not those of European descent. A strong genetic association with early-onset disease was supported by the documentation of numerous NRAMP1 alleles in children, whereas only a few were recovered from adult patients.

A recent GWAS conducted among populations from Gambia and Ghana identified a “gene desert” on chromosome 18 as a risk factor for pulmonary TB and a second locus on chromosome 11p13 as protective against TB. These associations were not consistent when repeated in other populations, thus suggesting that GWAS may have a limited impact on predisposition to adult pulmonary TB, at least when considered as a single phenotype.

### Conclusion

The exact role of genetic factors involved in TB remains unknown. Genetic heterogeneity along with a complex mode of inheritance and other factors such as the intensity

of exposure to TB and variable *Mycobacterium tuberculosis* strain virulence may also alter the clinical course of TB [23–25].

### 23.10.2 *Mycobacterium leprae*

Leprosy, also known as Hansen’s disease, is a chronic infectious disease caused by *Mycobacterium leprae* in which susceptibility to the mycobacteria and its clinical manifestations are affected by the host immune response. Leprosy patients can be classified into one of the five groups of the clinical spectrum depending on the stage in disease and host cell-mediated immunity. At one end of the spectrum is tuberculoid leprosy, paucibacillary, or designated TT, characterized by a small number of bacilli, a few skin lesions with well-defined edges, asymmetric neural involvement, acid fast bacilli absent or in small numbers in the skin and nerves. The local production of cytokines in this pole of the disease is type 1, including IFN- $\gamma$ , interleukins (IL-2, IL-7, IL-12, IL-15, and IL-18) and specific cellular immune response characteristics are preserved. At the other end is the lepromatous pole, also known as LL or multibacillary leprosy, characterized by large burden of bacilli in the nerves and skin, many lesions extensively infiltrated with macrophages and mycobacteria spongy tissue easily detected within this and nerves. A dysfunction of type-2 cytokine producing cell-mediated immunity (IL-4, IL-5, IL-10) occurs, resulting in anergic to *M. leprae* and high titers of circulating antibody. These patients have high potential to spread the disease.

It is suggested that human genetic factors may affect the acquisition of leprosy and the clinical course of disease. SNP-association studies showed a low lymphotoxin- $\alpha$  (*LTA*)-producing allele as a major risk factor for early onset leprosy.

A frequently occurring SNP in *TLR1* (the 602S allele), which impairs receptor trafficking and function, has been described and seems to play a protective role in the context of clinical leprosy. Previous studies have demonstrated that TLR2 mediates the innate immune recognition of *M. leprae*. TLR2 polymorphisms are associated with susceptibility to leprosy and/or leprosy reactions. The first genome-wide association study published for leprosy showed a significant association between polymorphisms found in seven genes and susceptibility to disease. The association between SNPs in *CCDC122*, *C13orf31*, *NOD2*, *TNFSF15*, *HLA-DR*, *LRRK2*, and *RIPK2* were stronger for MB leprosy. One attractive mechanism pointed out in the study links the nucleotide binding oligomerization domain-2 (NOD2)-mediated signaling pathway to susceptibility to infection with *M. leprae*. One additional study from Berrington et al. also suggests that NOD2 genetic variants are associated with susceptibility to leprosy and the development of reversal reaction and erythema nodosum leprosum [26,27].

### 23.10.3 *Chlamydia trachomatis*

Genetic polymorphism have also shown to affect clinical outcomes resulting from infections due to intracellular bacteria and other fastidious organisms such as *Chlamydia* spp. The genus *Chlamydia* comprises three important human disease-producing organisms. *C. trachomatis* causes blinding trachoma and sexually transmitted infections, whereas *Chlamydia pneumoniae* is associated with asthma and community-acquired pneumonia (CAP) and *Chlamydia psittaci* may cause severe respiratory systemic zoonotic infections.

Host genetic markers are the most promising biological predictors of complicated chlamydia infection at present. The importance of studying host genetic markers as well as behavioral markers linked to acquiring

chlamydia infection has been shown by research on *C. trachomatis* strains that cause trachoma. Bailey et al. found that up to 40% of the host-response to chlamydia infection in Gambian twin pairs is due to host genetics. They estimated the relative contribution of host genetics within the total variation in lymphoproliferative responses (specific T-cell immune responses) to *C. trachomatis* antigen. The scarring of the cornea (trachoma) and the scarring of the fallopian tubes (sexually transmitted chlamydia infection) have immunogenetic similarity. This has been summarized in a recent review which documented identical SNPs in ocular and tubal scarring for such as IL-10, TNF- $\alpha$ , and HLA types. In addition, MBL gene polymorphism has previously been described for ocular scarring.

Recent studies explored the possibility of genetic traits (carrying multiple SNPs in different genes) in the bacterial sensing system being associated with an aberrant immune response and subsequently with tubal pathology following a *C. trachomatis* infection. In one of those studies the authors assessed in subfertile women the presence of five SNPs in five genes, all encoding for pattern recognition receptors involved in sensing bacterial components. The SNPs were selected based on functional consequences they had for these genes; for instance, the genetic variation in the NOD2 gene resulted in a shorter protein due to a stop codon introduced by the SNP. It was shown that subfertile women with serological evidence of prior chlamydia infection (IgG) and two or more of these SNPs had a significantly higher chance of developing pathology than women with fewer or no SNPs. This study shows the importance of host genetic markers as indicators of late complications from chlamydia infection in women. This type of biomarker could be applied in better triage of women screened by the gynecologist for subfertility and laparoscopy [23,28].

#### 23.10.4 *Chlamydia pneumoniae*

*C. pneumoniae* may result in a wide spectrum of acute respiratory conditions such as CAP, bronchitis, and pharyngitis. Recent evidence also suggests its association with chronic conditions such as chronic obstructive pulmonary disease, cystic fibrosis (CF), and asthma exacerbations, thus raising the question of underlying host genetic susceptibility.

Recent in vitro studies proposed that underlying genetic factors such as IL-6-174G/C or low mannose-binding (MBL) polymorphisms are responsible for the development of chronic pulmonary conditions triggered by *C. pneumoniae*. Hence, host factors may determine the nature of *C. pneumoniae* infections (acute vs chronic) [23,29,30].

#### 23.10.5 *Mycoplasma pneumoniae*

*Mycoplasma pneumoniae* infections present as a wide spectrum of clinical pathology, including mainly conjunctivitis and respiratory diseases. It is the most prevalent agent of CAP in children above 2 years of age and has also been associated with asthma exacerbations, recurrent wheezing episodes, and acute bronchitis in children. More recent evidence demonstrates its role in CF exacerbations. In addition, *M. pneumoniae* has also been associated with immune-mediated conditions, resulting in erythema multiforma and reactive arthritis.

In vitro studies have demonstrated that *M. pneumoniae*-infected epithelial cells resulted in the induction of various cytokines and chemokines, including proinflammatory (TNF- $\alpha$ ), type 1 (IFN- $\gamma$ ) and type 2 (IL-6) cytokines, and  $\alpha$  (IL-8) and  $\beta$ -chemokines. In summary a predominant type 2-like cytokine response results from *M. pneumoniae* infections, as shown by various experimental studies. Thus individuals with stronger cytokine and cell-mediated immune responses may experience severe pulmonary injury, thus supporting the contribution of genetic polymorphisms [23].

#### 23.10.6 *Coxiella burnetii*

Another agent of atypical pneumonia is *Coxiella burnetii*, for which there is more evidence of the importance of underlying host polymorphisms. *C. burnetii*, an obligate intracellular bacterium, is the causative pathogen of Q fever. It is prevalent worldwide and highly infectious, being mainly transmitted by inhalation of contaminated aerosols. Infection generally results in acute Q fever, which is symptomatic in less than 40% of cases. A small proportion (less than 5%) will present with severe disease such as atypical pneumonia or granulomatous hepatitis. Less frequently, pericarditis, meningoencephalitis, or arthritis may also occur. Chronic Q fever, mainly endocarditis, observed among individuals with underlying predispositions (pregnancy, immunosuppression, or valvular



defects), occurs months to years after the acute episode and can result in significant mortality rates (25%–60%). Autoimmune conditions such as Libman–Sacks endocarditis have been reported in patients with documented *C. burnetii* infections. The highest rates of symptomatic Q fever are observed among men and children above 15 years of age, suggesting a major role for human genetic factors.

Immune control of *C. burnetii* results in granuloma formation and systemic cell-mediated immune responses, including IFN- $\gamma$ -production, probably as a result of monocyte/macrophage migration and T lymphocyte recruitment [23].

### 23.10.7 *Tropheryma whipplei*

*Tropheryma whipplei* is also an important agent of blood culture–negative endocarditis like *Coxiella*. *T. whipplei* is generally acquired through feco-oral transmission. While *T. whipplei* is found ubiquitously in the environment, it remains a rare disease with an annual incidence below 1 per 1000,000 population, suggesting the role of host factors. A higher prevalence of the bacteria is reported in men of European ancestry. Also, *T. whipplei* can be detected in as many as 35% of healthy carriers, thus reinforcing the role of an underlying polymorphism in determining late disease onset. The contribution of a genetic background to the evolution of Whipple disease was recently supported by the association of Whipple disease with HLA alleles DRB1\*13 and DQB1\*06. A genetic polymorphism in the cytokine genes was supported by the documentation of low levels of TGF- $\beta$ 1 and high production of IL-10, resulting in decreased Th1 and Th17 reactivity among infected patients compared with healthy controls, although the association was not significant. Polymorphisms in the cytokine genes should be further explored in larger cohorts, ideally by exome sequencing [23,31].

## 23.11 Host genetic susceptibility and protection from fungal infections

Much of our understanding of the molecular mechanisms underlying antifungal immunity was discovered by first investigating primary immunodeficiencies. This unique set of genetic deficiencies confers a predisposition to a narrow range of fungal pathogens and has often been the presenting clinical manifestation of the immunodeficiency itself.

### 23.11.1 *Candida*

*Candida* species are commensal organisms on the skin and the mucous membranes of healthy individuals but can become an opportunistic pathogen in the setting of a compromised immune system. *Candida* species are now considered a common pathogen in bloodstream infections. Traditional risk factors for invasive infection include recent treatment with broad-spectrum antibiotics, administration of parenteral nutrition, presence of intravascular catheters, prolonged intensive care unit stay, and neutropenia. Despite the aforementioned risk factors, only a minority of the population will develop an invasive fungal infection (IFI) in these circumstances. The causative etiology for developing an IFI is likely multifactorial; however, several monogenic diseases have been previously described to be associated with an increased susceptibility to infections with *Candida* species.

A notable example of a monogenic disorder, as described by Glocker et al., involved the identification of a homozygous point mutation in caspase recruitment domain–containing protein 9 in a consanguineous family with known recurrent IFIs. The mutated sequence was associated with a mean proportional reduction in Th17 cells as well as severe defects in dectin-1-triggered TNF- $\alpha$  signaling. The activation of this conserved pathway signifies its importance in fungal recognition, in stimulation of proinflammatory responses, and in Th17-cell differentiation.

Chronic mucocutaneous candidiasis (CMC) represents another classic immunodeficiency associated with recurrent fungal infections related to a defect in IL-17 and IL-22 immunity that is required for mucocutaneous antifungal host defense.

Autosomal recessive autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED) is a rare syndrome caused by a mutation in a thymic transcriptional regulator, AIRE, and manifests clinically in the form of impaired T-cell tolerance, self-reactive autoantibodies, and CMC. APECED patients form neutralizing antibodies against important antifungal cytokines including IL-17E, IL-17F, and IL-22. Furthermore, Liu et al. investigated the molecular mechanisms underlying STAT1 gain-of-function mutations known to characterize the AD form of CMC. These mutations increase STAT1-dependent responses through its impaired nuclear

dephosphorylation, influencing cellular response to multiple cytokines and culminating in the curbed differentiation of IL-17-producing T cells. In a follow-up study, Zheng et al. also reported that the negative effects of gain-of-function STAT1 mutations on STAT3 function were secondary to decreased histone acetylation, which affected STAT3 promoter binding and gene expression. These investigators reported that the reduction in STAT3 gene expression explains the observed low Th17 responses that are the hallmark of the clinical syndrome of CMC. They noted a reversal in the cellular transcriptional profiles by inhibiting STAT1 activation or by enhancing histone, suggesting that targeting epigenetic modifiers may offer a unique treatment strategy for AD CMC.

Hyper-IgE syndrome (HIES) is the result of a heterogeneous group of missense mutations or in-frame deletions most commonly in the STAT3 transcription factor that clinically manifests as a distinct syndrome characterized by extreme elevations in IgE levels, eczema, recurrent pulmonary infections, invasive aspergillosis (IA), and cold staphylococcal skin abscesses. Mutations in DOK8 and Tyk 2 have also been associated with HIES.

In addition to CMC, these patients are known to develop IFIs with filamentous molds including *Aspergillus* and *Scedosporium*. Pneumatocoles, parenchymal lung damage, and bronchiectasis resulting from a history of recurrent bacterial infections serve as a medium for these opportunistic mold infections [32,33].

### 23.11.2 Aspergillosis

Aspergillosis presents as a wide spectrum of diseases caused by fungi of the genus *Aspergillus* with clinical manifestations ranging from colonization (e.g., aspergilloma), to allergic bronchopulmonary aspergillosis, to disseminated forms of infection. IA has been estimated to occur in 10% of acute myeloid leukemia patients during postinduction aplasia or consolidation therapy and after 5%–15% of allogeneic hematopoietic stem cell transplants (HSCT). Other individuals at risk for IA include recipients of solid organ transplants and patients with chronic granulomatous disease (CGD).

#### 23.11.2.1 Genetic variability of host and susceptibility to invasive aspergillosis

The inborn deficiency of the phagocyte nicotinamide adenine dinucleotide phosphate oxidase leading to CGD is a well known example of primary immunodeficiency with predisposition to IA. Because of the impaired production of reactive oxygen species, patients with CGD often develop IA, typically within the first decade of life. Interestingly, these patients are uniquely susceptible to diseases with the *Aspergillus nidulans* complex, which are less virulent molds that seldom cause infection in immunocompromised patients. Keeping in mind the crucial requirement of innate immunity for effective antifungal host defense, several studies have revealed associations between genetic variants in components of the innate immune system and risk for IA. One classic example is a donor haplotype in TLR4 reported to increase susceptibility to infection after HSCT, especially if combined with cytomegalovirus seropositivity. In addition, and given the pivotal role of dectin-1 in fungal sensing, it is also not surprising that human dectin-1 deficiency has been reported to contribute to susceptibility to IA.

Of interest, genetic and functional deficiency of other molecules with opsonic activity such as, MBL and PTX3 have also been shown as major determinants of susceptibility to IA, pointing to a leading contribution of humoral immunity in response to *Aspergillus*.

#### Conclusion

The identification of patient-specific prognostic signatures of susceptibility to IA in high-risk patients is currently a major priority in the fields of hematology and microbiology. Ultimately, the discovery of reliable markers of susceptibility consistently associated with risk of IA, and functionally correlated with impaired antifungal

mechanisms of the host, may be a turning point toward innovative stratification strategies based on genetic screening or immune profiling, with the aim of predicting risk and severity of disease, efficacy of antifungal prophylaxis and therapy. This could eventually contribute to the successful design of antifungal vaccines [34].

### 23.11.3 *Cryptococcus neoformans* and *Cryptococcus gattii*

The predominant etiological agents of cryptococcosis are fungal pathogens that cause disease ranging from a mild pneumonia to life-threatening infections of the central nervous system. Resolution or worsening of *Cryptococcus* infection is determined following compound interactions of several host and pathogen-derived factors. The virulence of the organism relies on specific molecular factors that target host defenses to allow establishment of infection.

*Cryptococcus neoformans* is considered an opportunistic fungal pathogen and is the leading cause of fungal meningitis in the world. This infection remains the fourth leading cause of death in HIV-infected patients for which the burden of *Cryptococcus* infection remains high. An estimated 1 million cases of cryptococcal meningitis occur annually in these patients worldwide. Other risk factors for cryptococcosis include solid organ transplantation, hematological malignancy, and prolonged immunosuppression or chemotherapy. Multiple studies have demonstrated an increased propensity for cryptococcal infection in patients with CD4 lymphopenia. Defective production of IFN- $\gamma$  and TNF- $\alpha$ , two cytokines known to be involved in stimulation of anticryptococcal mechanisms in phagocytes, suggests that the deficiency of these proinflammatory mediators plays a crucial role in the development of IFIs in these patients. This concept was further supported when recombinant IFN- $\gamma$  was administered as an adjunctive therapy to one of the progressively deteriorating patients that led to clinical recovery. In addition, anti-IFN- $\gamma$  antibodies have been isolated from a series of patients with cryptococcosis and even the presence of autoantibodies to GM-CSF has been linked to cryptococcal infections. Recent investigations into the genetic susceptibility of cryptococcosis have begun to identify potential human genes associated with cryptococcosis. There has also been an association with genetic variability in the MBL complex and cryptococcosis, and we are starting to see reports of affected patients with mutations in cytokine genes for adaptive immunity such as IL-12. Thus both by phenotype and genotype, the search is underway for a detailed understanding of the genetic susceptibility to cryptococcosis, especially in the apparently immunocompetent host [35].

## 23.12 Malaria

The World Health Organization (WHO) estimated 225 million malaria cases worldwide with 781,000 deaths due to *Plasmodium* infection every year. Four species of *Plasmodium* (*Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, and *Plasmodium ovale*) are responsible for almost all human infections.

Malaria has been associated to gene selective pressure in the human genome, and it has been associated as an evolutionary force of some genetic diseases. A protective effect against malaria infection has been associated with genetic disorders involving various components of red blood cell (RBC), such as cytoskeleton, surface antigen, enzymatic machinery, or hemoglobin alterations [36].

### 23.12.1 Membrane and enzymatic disorders of red blood cells

Hereditary spherocytosis is a disorder characterized by the abnormal membrane lipid surface. This common hemolytic anemia reflects ineffective integral protein interactions and is associated with lower parasitemia. Other RBC membrane inherited disorders are hereditary ovalocytosis, elliptocytosis, pyropoikilocytosis, and acanthocytosis. Elliptocytosis has demonstrated resistance against invasion by *P. falciparum* in humans. Ovalocytosis is an RBC-inherited cytoskeleton disorder. South Asian ovalocytosis, also known as, Melanesian elliptocytosis have been associated with resistance to malaria infection, particularly against *P. falciparum* merozoites invasion. The Duffy, also called Duffy antigen/chemokine receptor (DARC), Fy glycoprotein, or CD234, is an RBC antigen encoded by the DARC human gene that is a nonspecific receptor for several chemokines [37]. The *P. vivax* merozoite utilizes the Fy antigen to invade RBC. The Fy antigen possesses two distinct alleles known as Fya and Fyb. Erythrocytes expressing Fya showed 41%–50% lower binding to *P. vivax* compared with Fyb, and individuals with the Fya + b-phenotype showed a 30%–80% lower risk of developing clinical symptoms of vivax malaria.

The G6PD deficiency and low levels of pyruvate kinase are the most commonly seen genetic alterations in RBC that can affect malaria outcomes. Infected erythrocytes deficient of G6PD were more phagocytosed by monocytes, which might be associated with the reduction of the parasitic load of the disease. PK deficiency is the second most common cause of hereditary nonspherocytic hemolytic anemia in humans. PK catalyzes the rate-limiting step of glycolysis, and the energy for erythrocytes is acquired from glycolysis, as RBC lack mitochondria. Its deficiency has been associated with the reduced survival and increased phagocytosis of parasite-infected erythrocytes [38].

#### 23.12.1.1 Hemoglobin alterations—hemoglobinopathies

Hemoglobinopathies are inherited disorders of Hb that can be classified into two major groups, that is,

1. structural alterations or variants of Hb, such as HbS, HbC, HbE; and
2. synthesis defects of Hb with a decrease or absence of globin chain synthesis such as alpha and beta-globin chains (alpha- and beta-thalassemia, resp.).

Several subsequent studies reported that  $\alpha^+$ -thalassemia (impaired but not absent synthesis of alpha-globin) was associated with a reduced risk of uncomplicated malaria episodes or a protective effect against severe forms of malaria. In beta-thalassemia the heterozygote of this inherited trait is associated with a mild anemia and an ineffective erythropoiesis, while the homozygote mutant is associated with severe anemia and the risk of early death. The beta-thalassemia trait is associated with a relative resistance against *P. falciparum* malaria and protection against severe form of malaria.

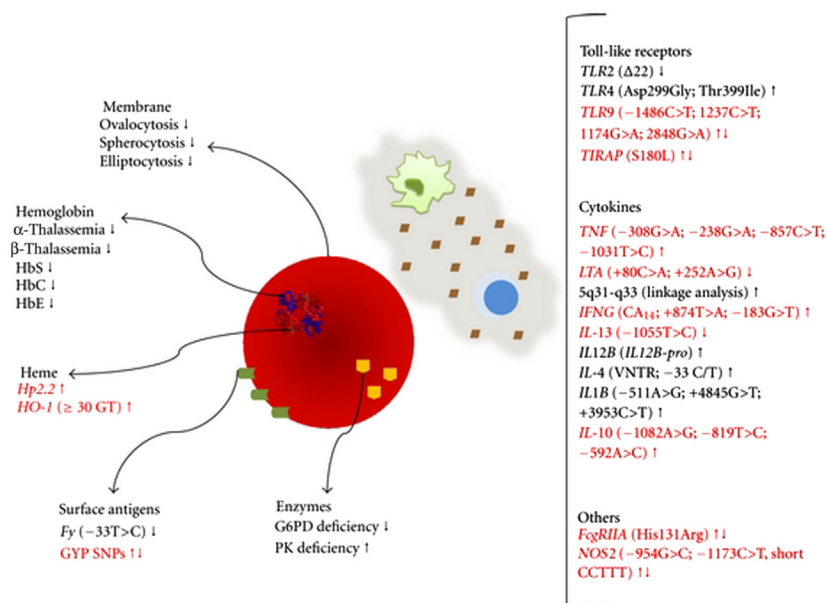
HbS polymerizes under deoxygenate conditions, and parasites become severely affected; HbS red blood cells become sickled, with the increased phagocytosis of infected erythrocytes. The presence of HbS in severe malaria patients is associated with decreased hemolysis and lower levels of free heme. Many studies have described an association between the heterozygote HbAS and protection against malaria, with more than 90% protection against severe forms. The red blood cells from HbE homozygous individuals are microcytic, with low hemoglobin concentrations, which reduces the ability for merozoite invasion and impairs parasite growth within this variant RBC. The presence of both protective factors, HbC and HbE, has been associated with a lower risk of developing severe forms of malaria.

### 23.12.1.2 Systemic regulation of heme

Heme reduces the production of prostaglandin E2 (PGE2) and TGF- $\beta$  from mononuclear cells through superoxide dismutase-1 (SOD-1), an enzyme responsible for the detoxification of harmful superoxide. SOD-1 is a powerful predictor of malaria severity in individuals infected with *P. vivax* with higher sensitivity and sensibility than TNF- $\alpha$  levels, confirming the importance of this enzyme in malaria pathogenesis. Studies have shown an association between several SNPs in the SOD-1 gene and different expressions of this enzyme in subjects with malaria [39].

### 23.12.2 Immune response

TNF- $\alpha$  is a proinflammatory cytokine that has attracted interest because of its ambiguous activity in host defense and pathogenesis of cerebral malaria and other serious complications. High concentrations of TNF- $\alpha$  are related to the development of symptoms associated with malaria, such as fever, and severe forms of infection, such as cerebral malaria. IL-12 is a proinflammatory cytokine that augments erythropoietic responses in infections with *Plasmodium* parasites. Lower levels of IL-12 have been associated with the pathogenesis of malaria in children and nonimmune adults through the augmentation of IFN- $\gamma$  release from cells of the innate immune system, while high levels of this cytokine are associated with severe malaria [36] (Fig. 23.1).



**FIGURE 23.1** Influence of erythrocyte and immune response gene polymorphisms in malaria outcome. The diagram summarizes the major genetic alterations identified in the erythrocyte and immune response pathways that influence malaria outcome. The up arrow indicates susceptibility, and the down arrow indicates resistance to malaria. Contradictory or not confirmed results are represented by red font color.

### 23.12.3 Malaria vaccine

Malaria vaccine was developed by assessing many new vaccine technologies including novel adjuvants, vectored prime-boost regimes, and the concept of community vaccination to block malaria transmission. Of all these, RTS,S is the first and, to date, the only vaccine to show a protective effect against malaria and provide meaningful public health benefit by reducing the burden of malaria when used alongside currently available interventions such as bed nets and insecticides among young. RTS,S is a recombinant protein-based malaria vaccine. Phase III results showed that three doses of RTS,S reduced clinical malaria by approximately half in children 5–17 months of age at first vaccination [40]. In a subsequent analysis after 18 months of follow-up, children vaccinated with RTS,S experienced 46% fewer cases of clinical malaria, compared to children immunized with a comparator vaccine. Hence, The WHO Regional Office for Africa announced on April 24, 2017, that Ghana, Kenya, and Malawi will partner with WHO in the Malaria Vaccine Implementation Programme that will make the RTS,S vaccine available in selected areas of the three countries, beginning in 2018.

## 23.13 Genetics of susceptibility to enteral pathogens

Variants in genes that encode molecules that modulate attachment, pathogen recognition, inflammatory cytokine response, innate, and acquired immunity are being identified as determinants of host genetic susceptibility to infectious diarrhea [41].

### 23.13.1 Host receptors used by enteral pathogens and their role in susceptibility

Enteric pathogens frequently use host molecules as their specific receptors. Such is the case of histo-blood group antigens (HBGAs) that are carbohydrates that contain structurally related saccharide molecules. The fucose transfer of ABH antigens in erythrocytes is catalyzed by FUT1, a member of the fucosyltransferase family, whereas FUT2 catalyzes a different fucosyltransferase in saliva and mucosal secretions. Human HBGAs may serve as receptors for numerous pathogens, including *Campylobacter jejuni*, *Helicobacter pylori*, and norovirus that bind to Se<sup>+</sup> cells and *Escherichia coli* and *Staphylococcus aureus* that bind to Se<sup>−</sup> cells. Norovirus GI-1 (Norwalk virus) binds preferably to O Se<sup>+</sup> cells, while G-II-3 and GII-4, the worldwide predominant epidemic strains, binds preferably to A Se<sup>+</sup> cells, Se<sup>−</sup> individuals are genetically immune to those individual norovirus infections. *E. coli* heat-labile (hLT) and cholera (CT) toxins are also able to use the HBGAs as receptors. hLT binds to both A and B antigens. Since the H antigen (O group) lacks this terminal sugar residue, the toxin is unable to bind the unmodified H antigen (O group) at all [7]. The toxins hLT and CT demonstrate distinct differences in affinity for the HBG, and interestingly among *Vibrio cholerae*, there are also biotype and serogroup specific differences, *V. cholerae* O1 demonstrating the highest affinity. Individuals with blood group O are 50% less likely to become infected with *V. cholerae* than non-blood group O individuals.

### 23.13.2 Innate immune genes associated with increased susceptibility to enteral pathogens

The NOD2 modulates the signaling induced by TLR4. The function of some TLRs, especially TLR4, may also depend on other associated cell surface coreceptors, such as the CD14 and lymphocyte antigen 96 (LY96, also known as MD-2). There is clinical evidence demonstrating that adults from developed countries possessing the CC genotype at the −4191 position of the CD14 promoter gene, associated with lower sCD14 expression, have a 36% higher risk of developing infectious diarrhea when traveling to a developing country. This observation is more significant for individuals developing diarrhea due to invasive organisms.

### 23.13.3 Innate immune response and cellular injury

The activated epithelial cells can synthesize and release IL-8 when exposed to enteropathogens. IL-8 stimulates the recruitment and transmigration of neutrophils into the intestinal lumen. The production of IL-8 in the intestine is genetically determined by at least one functional polymorphism. Individuals homozygous for the AA genotype at the −251 position of the IL-8 gene promoter produce significantly greater concentrations of IL-8 when stimulated and are at an increased risk of diarrhea due to enteroaggregative *Escherichia coli* (EAEC) and



*Clostridium difficile*-associated disease (CDAD) [42]. Symptomatic CDAD subjects have a 3.3 increase in odds of having the −251 AA SNP when compared with matched controls.

### 23.13.3.1 Acquired immunity

The cellular immune activation depends on the antigen processing and presentation to the CD4 + T cells via the HLAs, specifically MHC class II. Once a T cell recognizes a peptide within an MHC class II molecule, it can activate B cells. Different MHC class II alleles can affect the presentation of specific antigens and alter the resistance to different enteric infections.

A study in two Vietnamese populations reported a genetic association between typhoid fever and genes in the MHC class II. The HLA-DRB1 \*0301/6/8 and HLA-DQB1 \*0201-3 antigens were found more frequently in individuals with typhoid fever compared with healthy controls, whereas the HLA-DRB1 \*04, \*1001, and HLA-DQB1 \*0401/2 antigens were significantly more frequently found in healthy subjects [43].

A cohort study in Bangladesh found a protective association of the HLA class II allele DQB1 \*0601 and the heterozygous haplotype DQB1 \*0601/DRB1 \*1501 with *Entameba histolytica* infection [44]. Another cohort study done in Bangladeshi children found an association between HLA-DQB1 \*0301 and asymptomatic infection with *Cryptosporidium* spp. Of notice, the authors also found an association with the HLA class I B\*15 allele, suggesting that in addition to traditional helper T cells, other components of the cellular immune response are involved with parasite eradication.

## 23.14 Conclusion

Although there has been tremendous progress in molecular microbiology testing, challenges remain. Currently, majority of clinical laboratories do not perform molecular microbiology testing. To increase access to these important tests, simple platforms with a broad test menu are needed. There are molecular tests for the identification of some bacterial pathogens directly from a clinical specimen, most notably methicillin-resistant *S. aureus*. Moreover, simple sequencing methods and comprehensive databases are now available and are routinely used in some referral and large clinical laboratories for the identification of bacteria, mycobacteria, fungi, and parasites that are not easily identified using conventional methods. Currently, molecular methods have not replaced standard culture methods for routine bacterial identification and susceptibility testing, and this is viewed by some as the next big challenge. For molecular testing to replace the currently used culture-based methods, and for us to move from phenotypic to genotypic methods for susceptibility testing, advances in technology, as well as a better understanding of the genetics of pathogen resistance, are needed. Although there are skeptics who feel this is not possible, with the advances in molecular microbiology that have occurred in the past decade, the potential for the next decade appears limitless.

## References

- [1] Sobel M, Bagg A, Caliendo AM, Ladanyi M, Zehnbauser B. The evolution of molecular genetic pathology. *J Mol Diagn* 2008;10(6):480–3.
- [2] Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;6:95–108.
- [3] Chatterjee K. Host genetic factors in susceptibility to HIV-1 infection and progression to AIDS. *J Genet* 2010;89:109–16.
- [4] Singh KK, Spector SA. Host genetic determinants of HIV infection and disease progression in children. *Pediatr Res* 2009;65:55–63.
- [5] Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996;382:722–5.
- [6] Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 2005;307:1434–40.
- [7] Holmner A, Askarieh G, Okvist M, Krengel U. Blood group antigen recognition by *Escherichia coli* heat-labile enterotoxin. *J Mol Biol* 2007;371:754–64.
- [8] Wang L, Zou Z, Wang K. Clinical relevance of HLA gene variants in HBV infection. *J Immunol Res* 2016;2016:9069375.
- [9] Li S, Qian J, Yang Y, et al. GWAS identifies novel susceptibility loci on 6p21.32 and 21q21.3 for hepatocellular carcinoma in chronic hepatitis B virus carriers. *PLoS Genet* 2012;8(7):e1002791.
- [10] Bengsch B, Thimme R, Blum H. Role of host genetic factors in the outcome of hepatitis C virus infection. *Viruses* 2009;1(2):104–25.
- [11] Rehmann B. Hepatitis C virus versus innate and adaptive immune responses: a tale of coevolution and coexistence. *J Clin Invest* 2009;119:1745–54.
- [12] Götte M, Feld JJ. Direct-acting antiviral agents for hepatitis C: structural and mechanistic insights. *Nat Rev Gastroenterol Hepatol* 2016;13:338–51.

- [13] Chattopadhyay K. A comprehensive review on host genetic susceptibility to human papillomavirus infection and progression to cervical cancer. *Indian J Hum Genet* 2011;17(3):132–44.
- [14] Bubenik J. Genetically modified cellular vaccines for therapy of human papilloma virus type 16 (HPV 16)-associated tumours. *Curr Cancer Drug Targets* 2008;8(3):180–6.
- [15] Houldcroft CJ, Kellam P. Host genetics of Epstein–Barr virus infection, latency and disease. *Rev Med Virol* 2015;25(2):71–84.
- [16] Booth C, Gilmour KC, Veys P, et al. X-linked lymphoproliferative disease due to SAP/SH2D1A deficiency: a multicenter study on the manifestations, management and outcome of the disease. *Blood* 2011;117:53–62.
- [17] Zhizhuo H, Junmei X, Yuelin S, et al. Screening the PRF1, UNC13D, STX11, SH2D1A, XIAP, and ITK gene mutations in Chinese children with Epstein–Barr virus-associated hemophagocytic lymphohistiocytosis. *Pediatr Blood Cancer* 2012;58:410–14.
- [18] Phuong Lan NT, Hirayama K. Host genetic susceptibility to severe dengue infection. *Trop Med Health* 2011;39(4 Suppl):73–81.
- [19] Pang T, Mak TK, Gubler DJ. Prevention and control of dengue—the light at the end of the tunnel. *Lancet* 2017;17(3):79–87.
- [20] Charbonnel N, Pagès M, Sironen T, Henttonen H, Vapalahti O, Mustonen J, et al. Immunogenetic factors affecting susceptibility of humans and rodents to hantaviruses and the clinical course of hantaviral disease in humans. *Viruses* 2014;6(5):2214–41.
- [21] Horby P, Nguyen NY, Dunstan SJ, Baillie JK. The role of host genetics in susceptibility to influenza: a systematic review. *PLoS One* 2012;7(3):e33180.
- [22] Gika AD, Rich P, Gupta S, Neilson DE, Clarke A. Recurrent acute necrotizing encephalopathy following influenza A in a genetically predisposed family. *Dev Med Child Neurol* 2009;52:99–102.
- [23] Asner SA, Morré SA, Bochud PY, Greub G. Host factors and genetic susceptibility to infections due to intracellular bacteria and fastidious organisms. *Clin Microbiol Infect* 2014;20(12):1246–53.
- [24] Moller M, de Wit E, Hoal EG. Past, present and future directions in human genetic susceptibility to tuberculosis. *FEMS Immunol Med Microbiol* 2010;58:3–26.
- [25] Bustamante J, Boisson-Dupuis S, Abel L, Casanova J. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN- $\gamma$  immunity. *Semin Immunol* 2014;26(6):454–70.
- [26] Alter A, Alcaïs A, Abel L, Schurr E. Leprosy as a genetic model for susceptibility to common infectious diseases. *Hum Genet* 2008;123:227–35.
- [27] Pinheiro RO, de Souza Salles J, Sarno EN, Pereira Sampaio E. *Mycobacterium leprae*–host-cell interactions and genetic determinants in leprosy: an overview. *Future Microbiol* 2011;6(2):217–30.
- [28] Morre SA, Karimi O, Ouburg S. *Chlamydia trachomatis*: identification of susceptibility markers for ocular and sexually transmitted infection by immunogenetics. *FEMS Immunol Med Microbiol* 2009;55:140–53.
- [29] Rantala A, Lajunen T, Juvonen R, et al. Interleukin-6-174G/C promoter polymorphism is associated with persistence of *Chlamydia pneumoniae* antibodies in young men. *Scand J Immunol* 2011;74:95–9.
- [30] Rantala A, Lajunen T, Juvonen R, et al. Low mannose-binding lectin levels and MBL2 gene polymorphisms associate with *Chlamydia pneumoniae* antibodies. *Innate Immun* 2011;17:35–40.
- [31] Schinnerling K, Moos V, Geelhaar A, et al. Regulatory T cells in patients with Whipple’s disease. *J Immunol* 2011;187:4061–7.
- [32] Lionakis MS. Genetic susceptibility to fungal infections in humans. *Curr Fungal Infect Rep* 2012;6(1):11–22.
- [33] Pana ZD, Farmaki E, Roilides E. Host genetics and opportunistic fungal infections. *Clin Microbiol Infect* 2014;20(12):1254–64.
- [34] Cunha C, Aversa F, Romani L, Carvalho A. Human genetic susceptibility to invasive aspergillosis. *PLoS Pathog* 2013;9(8):e1003434.
- [35] Rohatgi S, Pirofski LA. Host immunity to *Cryptococcus neoformans*. *Future Microbiol* 2015;10(4):565–81.
- [36] Kwiatkowski D, Hill AVS, Sambou I, et al. TNF concentration in fatal cerebral, non-fatal cerebral, and uncomplicated *Plasmodium falciparum* malaria. *Lancet* 1990;336(8725):1201–4.
- [37] Maestre A, Muskus C, Duque V, et al. Acquired antibody responses against *Plasmodium vivax* infection vary with host genotype for duffy antigen receptor for chemokines (DARC). *PLoS One* 2010;5(7):e11437.
- [38] Williams TN. Human red blood cell polymorphisms and malaria. *Curr Opin Microbiol* 2006;9(4):388–94.
- [39] Mendonça VR, Luza NF, Santosa NJG, et al. Association between the haptoglobin and heme oxygenase 1 genetic profiles and soluble CD163 in susceptibility to and severity of human malaria. *Infect Immun* 2012;80(4):1445–54.
- [40] Olotu A, Fegan G, Wambua J, Nyangweso G, Leach A, Lieven M, et al. Seven-year efficacy of RTS,S/AS01 malaria vaccine among young African children. *N Engl J Med* 2016;374:2519–29.
- [41] Flores J, Okhuysen PC. Genetics of susceptibility to infection with enteric pathogens. *Curr Opin Infect Dis* 2009;22(5):471–6.
- [42] Jiang ZD, DuPont HL, Garey K, et al. A common polymorphism in the interleukin 8 gene promoter is associated with *Clostridium difficile* diarrhea. *Am J Gastroenterol* 2006;101:1112–16.
- [43] Dunstan SJ, Stephens HA, Blackwell JM, et al. Genes of the class II and class III major histocompatibility complex are associated with typhoid fever in Vietnam. *J Infect Dis* 2001;183:261–8.
- [44] Duggal P, Haque R, Roy S, et al. Influence of human leukocyte antigen class II alleles on susceptibility to *Entamoeba histolytica* infection in Bangladeshi children. *J Infect Dis* 2004;189:520–6.



# Molecular mechanisms in cancer susceptibility—lessons from inherited cancers

Shirley V. Hodgson<sup>1,2</sup>

<sup>1</sup>Cancer Genetics, Department of Clinical Genetics, St. George's University of London, London, United Kingdom

<sup>2</sup>Cancer Genetics Consultant, Leicester Royal Infirmary, Leicester, United Kingdom

## 24.1 Introduction

The role of inherited characteristics in cancer susceptibility became appreciated in the last century through observing occasional families in which specific cancers appeared to be inherited in an autosomal-dominant fashion, and later with the observation that a family history of certain cancers in close relatives conferred an increased risk of the same cancer type, the risk increasing with increasing numbers of affected relatives.

This chapter provides a brief overview on molecular mechanisms in cancer pathogenesis based on knowledge and clinical experience derived from inherited cancer susceptibility and the cancer family syndromes.

## 24.2 Inherited and familial cancer

It soon became clear that there were some families in which certain individuals appeared to have a strong inherited susceptibility to specific cancers, such as breast and ovarian cancer, or colorectal and endometrial cancer. This susceptibility appeared to be inherited as an autosomal-dominant trait. In addition, the autosomal-dominant inherited form of adenomatous colorectal polyposis, which conferred a strong susceptibility to colorectal cancer (familial adenomatous polyposis, FAP) had been defined in the mid-1900s, and registries set up to ascertain and manage families in which this condition was segregating. This was because it was found to be helpful to see affected individuals and their families in specialist centers and arrange screening for their offspring and other at-risk relatives, to reduce their cancer risk.

The epidemiological evidence led to the concept that some of the observed familial clustering of cancers in the population could be due to the presence in the population of rare genes conferring a strong cancer susceptibility. However, this could not account for all of the observed familial clustering, so other more common genes conferring a lesser degree of cancer susceptibility were postulated also to be involved, interacting with environmental factors. This concept has been generally substantiated by the identification of rare high-penetrance genes in which mutations caused a strong susceptibility to specific cancers, such as *BRCA1* and *BRCA2*, which confer high risks of breast and ovarian cancer [relative risk (RR) in mutation carriers about 10–15], less rare genes in which mutations conferred a moderately increased risk of specific cancers, such as *CHEK2* for breast cancer (RR about 2), and common genetic variants conferring only slight alterations in cancer risk (RR about 1.1–1.3), which are being identified by genome-wide association studies (GWAS), and are thought to act multiplicatively to alter cancer susceptibility.

### 24.3 Oncogenes and tumor suppressor genes

The first breakthrough in our understanding of inherited cancer susceptibility came from the observation by Knudson that inherited (bilateral) forms of retinoblastoma, a rare cancer of the retina seen in childhood, developed at an earlier age than unilateral, sporadic cases. He demonstrated a different mathematical relationship between incidence and age at diagnosis between the sporadic and familial cases, and he interpreted this as indicating that initiation of the inherited form required a single molecular event (hit) in the retina at a susceptible age, whereas the sporadic form required two “hits” in the same retinal cell. This became known as the “two-hit hypothesis,” and Knudsen went on to postulate that the molecular “hit” in question was the inactivation of a tumor suppressor gene (TSG) in the susceptible tissue, and that tumor initiation required both TSGs to be inactivated in a single retinal cell in childhood. This theory has proved to be very prescient, since we now know that there are many autosomal-dominant conditions causing an inherited susceptibility to specific cancers, which are due to the inheritance of a faulty TSG. These include FAP, Gorlin syndrome, Cowden syndrome (CS), and Peutz–Jeghers syndrome (PJS) [1], many of which are characterized by specific clinical physical features (e.g., melanin freckles on the skin and mucous membranes in PJS). A form of hamartomatous polyposis known as juvenile polyposis can also be due to inherited abnormalities in TSGs, including *SMAD4* and *BMPRI1A* [2,3]. TSGs are known as “gatekeepers” and are usually involved in the initiating steps of cancer development; in FAP for instance, multiple adenomas develop in the colon in the teenage years, and these benign polyps require further additional mutations to occur within them for progression to cancer. Loss of both functional copies of the *APC* gene (which is the gene that, if mutated, causes FAP) in a single cell is often seen as an early initiating molecular event in colorectal cancers, although the process is somewhat more complex than this, and different colorectal cancers appear to develop along different molecular pathways. The *APC* gene has several functions, but one of these is to sequester the beta catenin gene for degradation; if not degraded, beta catenin stimulates the growth of colonic cells. This underlies the TSG function of the *APC* gene [4].

Another example of a condition caused by the inheritance of a faulty TSG is CS. This is an autosomal-dominant inherited disorder characterized by benign and malignant tumors, mainly thyroid and breast cancers, and usually caused by inherited mutations in the *PTEN* gene. *PTEN* is a major negative regulator of the phosphatidylinositol 3-kinase/protein kinase B/mammalian target of rapamycin signaling pathway, controlling growth, protein synthesis, and proliferation. In cases without *PTEN* mutations, inherited mutations in genes with related functions, such as *AKT1*, *PIK3CA*, *SDHB-D*, *KLLN* promoter methylation, and *SEC23B*, may be detected [5].

### 24.4 DNA repair genes

Another important cause of inherited cancer susceptibility is the inheritance of faulty DNA repair genes. There are several autosomal-recessive conditions known to confer an increased risk of specific cancers, which are caused by biallelic mutations in DNA repair genes, such as xeroderma pigmentosum, ataxia telangiectasia, and Fanconi anemia.

#### 24.4.1 The breast and ovarian cancer

Both the *BRCA1* and *BRCA2* breast/ovarian cancer susceptibility genes participate in the repair of double-strand breaks in DNA, and inherited mutations in these genes confer a strong susceptibility to breast and ovarian cancer. Mutations in other genes that participate in the same molecular DNA repair pathway, such as *ATM* and *CHEK2*, can confer a lesser degree of breast cancer susceptibility.

Breast cancer susceptibility is usually due to inherited alterations in genes involved in the DNA repair pathway for the repair of DNA mismatches. The most penetrant genes are *BRCA1* and *BRCA2*, which were identified by linkage analysis, and mutations in these genes confer a strong predisposition to breast cancer, and to a lesser extent to ovarian cancer [6]. The gene products participate in the repair of double-strand DNA breaks by error-free homologous recombination repair, preventing gross chromosome rearrangements. Moderate penetrance genes have more recently been identified by the candidate gene approach, and these also feature in the same DNA repair pathway, notably *ATM*, *CHEK2*, and *PALB2*. Some other high-penetrance genes are also occasionally implicated in causing breast cancer susceptibility, notably *CDH1*, *PTEN*, *STK11*, and *TP53*; but of these only the latter is mainly concerned with DNA repair, and the others are TSGs. The genomic instability in cancers arising



in individuals with *BRCA1* or *BRCA2* mutations makes the cancers more susceptible to treatment with agents that inhibit the repair of DNA breaks in a different molecular pathway, such as poly(ADP-ribose) polymerase (PARP) inhibitors. These seem to have an enhanced therapeutic effect in such individuals. Treatment with platinum salts is also more effective [7].

#### 24.4.2 Colorectal cancer

With regard to colorectal cancer, the autosomal-dominant condition known as Lynch syndrome, which confers high risks of colorectal and endometrial cancer and smaller increased risks of certain other extracolonic cancers, was found to be due to inherited mutations in one of several genes (*MLH1*, *MSH2*, *MSH6*, and *PMS2*), which repair DNA mismatches, recognize and repair replication errors, oxidative damage and alkylating damage to DNA [8–11]. In addition, autosomal-recessive adenomatous polyposis can be due to biallelic mutations in the *MutYH* and *NTHL1* genes, the normal function of which is to repair oxidative damage (*MutYH*) and base excision repair errors (*NTHL1*) in DNA [12]. Germline mutations in the proofreading polymerases *POLE* and *POLD1* cause autosomal-dominant polyposis and colorectal cancer susceptibility [13]. However, the commoner condition of autosomal-dominant FAP is caused by inherited mutations in a TSG, the *APC* gene [14].

The adenoma–carcinoma sequence in individuals with Lynch syndrome is much accelerated, thought to due to the fact that errors in microsatellite DNA in growth-promoting genes such as *TGFβ*, which remain unrepaired in the early neoplastic lesions of affected individuals, themselves promote growth. This has another effect, that of allowing the accumulation in the cancers of novel peptides that are immunogenic, so that individuals with Lynch syndrome appear to mount a strong immune response to their cancers (seen pathologically as enhanced lymphocytic infiltration of the tumors), which may result in improved survival rates from cancer in affected individuals. Tumors with DNA mismatch repair (MMR) deficiency tend to be resistant to treatment with 5-fluorouracil, but more sensitive to platinum agents, and immunomodulatory monoclonal antibodies are being developed to target immunosuppressive receptors on T lymphocytes in cancers in individuals with MMR deficiency [15]. Immunization to common novel peptides in microsatellite unstable tumors is also being considered.

Individuals with biallelic mutations in the MMR genes (*MSH6* and *PMS2*) have a condition known as constitutional MMR deficiency syndrome, with a high risk of childhood cancers including colorectal cancer, hematological malignancies, and brain tumors. Hyperpigmented skin lesions and pilomatricomas are a feature of this condition, with café-au-lait patches. A similar phenotype has been described in a child with a novel germline *POLE* mutation [16].

### 24.5 Cancer family syndromes

The inheritance of faulty oncogenes is a relatively uncommon cause of cancer susceptibility. There are several cancer susceptibility syndromes where neoplastic proliferation in multiple organs results from pathogenic mutations or polymorphic variants in an oncogene.

#### 24.5.1 Multiple endocrine neoplasia

The multiple endocrine neoplasia (MEN) type 2 (MEN2) is an excellent example, where inherited mutations in the *Ret* oncogene cause a susceptibility to medullary thyroid cancer, pheochromocytomas, and parathyroid adenomas. The risks of these tumors are high but vary with the site of the mutation in the gene [17]. Subtypes of MEN2 are MEN2A and MEN2B (now MEN type 3). The MEN type 1 (MEN1) is also an autosomal-dominant cancer family syndrome involving tumors in pituitary, parathyroid, and pancreas caused by mutations in *Menin* gene. A variant of MEN1 is named MEN type 4 caused by mutations in *CDKN1B*.

#### 24.5.2 RAS–MAPK syndromes

There are several childhood dysmorphic syndromes caused by low penetrance mutations in oncogenes in the RAS–MAPK pathway, known as the “RASopathies,” and there is a slightly increased risk of specific cancers in some of these conditions [e.g., leukemia in Noonan syndrome (NS)]. Interestingly, activating somatic mutations in the genes causing NS may be found in sporadic leukemias [18].

### 24.5.3 Von Hippel–Lindau disease and related syndromes

A different mechanism for the inheritance of susceptibility to specific cancers is seen in Von Hippel–Lindau (VHL) disease, an autosomal-dominant susceptibility to cerebellar hemangioblastomas, retinal angiomas, and renal cell carcinomas. These tumors are all very vascular, and the basic molecular defect is the failure to prevent the activation of the vascular growth-promoting HIF1 that stimulates vascular endothelial growth factor (VEGF), resulting in the overgrowth of vascular tissues. The *VHL* gene product plays a key role in cellular oxygen sensing by targeting and sequestering hypoxia inducible proteins such as the alpha subunit of the *HIF* gene product, preventing it from dimerizing with the beta subunit, which is needed for its normal function, and promoting its ubiquitination and degradation the proteasomes [19].

Genes with a similar effect of stimulating the VEGF pathway, if inherited in mutated form, are the *SDHB*, *SDHD*, and *SDHA*, and these are implicated in causing a susceptibility to vascular tumors of the autonomic nervous system such as paragangliomas, glomus tumors, and pheochromocytomas. A high proportion of such tumors, even when sporadic, occur in individuals with germline mutations in such genes. Inherited fumarate hydratase (*FH*) mutations cause a susceptibility to renal cancers and lipomatous skin lesions by a similar mechanism. Rarer genes occasionally found to cause pheochromocytomas and similar tumors include *KIF1Bbeta*, *MAX*, *SDHAF2*, and *TMEM127*. *NF1*, the gene that causes neurofibromatosis type 1 (NF1) when inherited in mutated form, is a TSG, but it can also cause pheochromocytomas [20].

### 24.5.4 Immunogenetics and cancer

There are other mechanisms involved in altering cancer susceptibility. Clearly the immune response to the foreign peptides associated with the growth of cancer is important in the body's defense against cancer, and we are beginning to appreciate that genes that regulate immune function do play a role in this. The main way in which this is being determined is by the finding in GWAS that certain immune regulating genes are associated with a small alteration in the risk of certain cancers.

### 24.5.5 RNA interference and cancer

A novel mechanism for inherited cancer susceptibility is the inheritance of a mutation in *DICER1*, a gene encoding a protein involved in the microRNA processing pathways. Some mutations in this gene appear to be oncogenic. Individuals with constitutional *DICER1* mutations are susceptible to cystic lung disease, nephromas, multinodular goiter, and Sertoli–Leydig tumors, among others, with a predilection for lung, kidney, ovary, and thyroid tissues [21]. Increasingly new molecular evidence is emerging on the role of RNA interference in the cancer pathogenesis.

---

## 24.6 Genetic imprinting and cancer

---

Imprinting is another way in which our genetic constitution can alter cancer susceptibility. There is good evidence that imprinting of certain genes is seen in cancers. Certain childhood overgrowth syndromes are due to imprinting of genes, such as Beckwith–Wiedemann syndrome, which may be caused by downregulation of (maternal) TSGs and upregulation of (paternal) growth-stimulating genes by imprinting. This causes a developmental overgrowth syndrome with an increased risk of Wilms tumor and other cancers. Mutations in these genes can also cause Beckwith–Wiedemann syndrome, but the cancer risk may be less in these cases [22].

The effect of environmental carcinogens is well known, as exemplified by the increased risk of lung cancer in smokers. However, the fact that some smokers get lung cancer and others do not may be related to their genetic ability to metabolize the carcinogens in tobacco smoke. GWAS are beginning to show that variants in genes which control the metabolism of certain carcinogens, are also involved in regulating susceptibility to a variety of different cancers [23].

A recent publication has indicated that breast cancers in multiple-case breast cancer families with no known genetic susceptibility mutations may in part be due to the contribution of heritable DNA methylation marks [including true epimutations and methylation quantitative trait loci (mQTLs)]. They identified 24 previously unknown epigenetic changes that alter a woman's risk of breast cancer and can be passed down through

generations without involving changes in the DNA sequences of genes. Further work is needed to substantiate this, but it indicates a novel way in which inherited cancer susceptibility might occur [24].

## 24.7 Complex cancer genomics

It is becoming clear that the genetic influences on cancer susceptibility can be divided into three broad groupings: rare mutations in genes, which confer a strong predisposition to specific cancers, such as *BRCA1* and *BRCA2* for breast/ovarian cancer, more common but still quite rare mutations in genes that confer a moderate cancer risk, and common variants in genes, which confer only a small alteration in risk. These latter probably act interactively with each other and environmental factors to contribute to cancer risk, and counseling for these and some moderate risk variants can be quite complex [25]. The functions of genes, which confer small increments in risk, as detected by GWAS, are also frequently involved in similar molecular pathways as the high-penetrance genes and can give insight into the molecular mechanisms of the specific cancer susceptibility.

Genome-wide scans for single-nucleotide polymorphisms (SNPs) are being employed in large populations of cancer cases and controls to ascertain SNPs that confer a very small increased risk of specific cancers. On their own these polymorphisms cannot usefully predict the susceptibility of an individual to a certain cancer, since many other genetic and environmental factors will also have a bearing on disease susceptibility, but it is possible that in the future, individuals carrying several high-risk variants will be identified, in whom the RR of a disease could be elevated sufficiently to warrant surveillance and preventative strategies to be offered. Thus individuals could be typed for a panel of SNPs, which individually may only confer an RR [odds ratio (OR)] of 1.2–1.3 but collectively could identify a small proportion of individuals who have a more substantial increase in risk [26,27]. In the case of colorectal cancer for instance, individuals with several such SNPs could have an overall RR of colorectal cancer >3, where regular colonoscopy screening might be offered. A similar situation may be encountered in breast cancer susceptibility, where a panel of common polymorphic loci identified in case–control studies could be tested, where the high-risk allele of each confers an OR of <1.2 on average. Women with six high-risk alleles, for instance, would have a significantly increased risk of breast cancer (OR about 2) and could be offered breast cancer screening earlier than that offered in the general population, and women with several low-risk alleles could be screened from a later age [26]. Such testing is now being introduced as part of a polygenic risk score that can be added to pedigree data (often assessed using the BOADICEA risk assessment for breast cancer) and will be clinically validated as research progresses.

Companies offering SNP risk profile services do not routinely offer pre- and posttest counseling and explanation of the results. Thus individuals informed that they possess a SNP conferring increased disease risk may wrongly perceive this risk to be substantial, and they may be unaware that the information does not take account of all other predictive factors (including other polymorphic loci, family history, mutations in high-penetrance genes, and lifestyle factors). Alternatively, one low-risk SNP result could lead to false reassurance and increased risk-taking behavior. Unregulated testing could lead to an increased workload for primary care practitioners, as patients may present requesting explanation of the results of the test, and access to further diagnostic testing. The implications of such multi-SNP tests are complex tests requiring sophisticated algorithms for interpretation, so difficult to interpret by most clinicians not working in the field of genetics.

## 24.8 Inherited susceptibility to leukemia

Leukemia affects approximately 1%–2% of the population in the West. B-cell chronic lymphocytic leukemia (CLL) is the most common, accounting for around 30% of all cases. There is an increased risk to the relatives of cases of leukemia, most marked in CLL. However, when familial clusters of leukemia are reported, the type of leukemia in individual relatives is not always concordant (Table 24.1).

Acute lymphatic leukemia (ALL) is the most common form in children, whereas acute myeloid leukemia (AML) predominates in adults. Genetic disorders that cause a predisposition to AML include Down syndrome (DS), Fanconi anemia, NF1, Kostmann syndrome, and DNA repair defects. Myeloproliferative disorders and AML, and monosomy 7 are the most common abnormality seen in familial cases of AML; familial clustering of AML with monosomy 7 may occur [28].

Inherited susceptibility to AML is rare, apart from in the conditions listed earlier. True nonsyndromic familial AML includes autosomal-recessive disorders with myelodysplasia and monosomy 7. An autosomal-dominant

**TABLE 24.1** Genetic disorders associated with leukemia

Ataxia telangiectasia
Autoimmune lymphoproliferative syndrome
Blackfan–Diamond syndrome
Bloom syndrome
Down syndrome
Dyskeratosis congenita
Fanconi anemia
Biallelic mismatch repair gene mutations
Immune deficiency diseases (e.g., severe combined immunodeficiency, common variable immunodeficiency)
Familial platelet disorder
Immunodeficiencies
Incontinentia pigmenti
Kostmann syndrome
Li–Fraumeni syndrome
Lynch syndrome
N syndrome
Neurofibromatosis type 1
Nijmegen breakage syndrome
Noonan, Cardiofaciocutaneous syndrome (CFC)
Rothmund–Thomson syndrome
Seckel syndrome
Shwachman syndrome
Trisomy 21
Wiskott–Aldrich syndrome

condition characterized by thrombocytopenia and platelet dense granule storage pool deficiency, platelet dysfunction, and a strong predisposition to AML and lymphoma, is well recognized. This is due to germline mutations in the *RUNX1* (runt-related transcription factor 1) (*AML1* and *CBFA2*) gene [29]. *RUNX1* was first identified as the gene rearranged in the somatic translocation t(8:21)(q22;q22.12) detected in patients with AML. The *RUNX1* protein complexes with the core-binding transcription factor (CBF) regulate genes involved in hematopoiesis. Familial AML can also be due to inactivating germline *CEBPA* (CCAAT-enhancer binding protein) mutations. This gene encodes the granulocyte differentiating factor C/EBP $\alpha$ . Homozygous or monoallelic germline inactivating mutations in *EZH2* (a histone methyltransferase) have been found in about 12% of patients with myelodysplastic and myeloproliferative disorders [30]. Conditions causing myelofibrosis also cause an increased risk of AML.

A rare form of AML, erythroleukemia (FAB-M6) may be familial (Di Guglielmo syndrome) in a minority of cases. Familial erythroleukemia is a leukemic or preleukemic state in which red cell proliferation occurs, with ineffective hyperplastic erythropoiesis and megaloblastic components accompanied by myeloblastic proliferation [31]. Mutations in the erythropoietin receptor may be causal.

Chronic myeloid leukemia (CML) is characterized by the Philadelphia (Ph) chromosome, present in over 90% patients, and some apparently Ph-negative cases of CML have a variant Ph translocation not cytogenetically apparent [32]. The Ph chromosome is also found in a small proportion of patients with ALL (20% adults, 5% children) and AML (2% adults). The classic Ph3 chromosome results from a translocation involving chromosomes

9 and 22 (t(9; 22)(q34; q11)). The breakpoint on chromosome 9 involves the Abelson oncogene (*ABL*), and on chromosome 22 the breakpoints occur within a small region designated as the breakpoint cluster region (*BCR*), a central segment in the *BCR* gene [33]. The 9;22 translocation results in the juxtaposition of proximal 5' *BCR* gene exons (1, 2, and 3) and *ABL* sequences (exons 2–11, exons 1a and 1b).

Most cases of CLL are B-cell type. This type of leukemia has the highest familial risk of all leukemias, with relatives of cases at significantly increased risk for CLL RR = 7.52 (3.63–15.56), for non-Hodgkin lymphoma RR 1.45 (0.98–2.16), and Hodgkin lymphoma RR 2.35 (1.08–5.08) [34,35].

An association between disordered immune function and lymphoreticular malignancy is suggested by the finding that in some reports of familial CLL, autoimmune or immunological disorders have been noted among unaffected relatives, and about 13% of apparently healthy relatives of patients with familial CLL have a monoclonal B lymphocyte population detectable by cell flow analysis (3% in controls). Rawstron et al. [36] detected subclinical levels of CLL-like cells in 14% of relatives compared to about 1% of normal controls suggesting that although the lifetime risk of CLL in familial cases is 20%–30%, there may be a significant incidence of subclinical disease.

Recent GWAS have identified about 22 loci that may contribute to small alterations in RR of leukemia, especially CLL [37]. These studies have found low penetrance risk alleles for CLL [38,39] and ALL, each conferring small RRs of disease (1.2–1.7 per allele), which act independently, such that the 2% of the population who carry 13 or more risk alleles would have an eightfold increased risk of disease. These variants are involved in lymphoid cell development [40].

Germline mutations in *DAPK1*, a  $\text{Ca}^{2+}$ /calmodulin-dependent serine/threonine kinase that acts as a positive regulator of apoptosis in part via phosphorylation of *p53* causing loss or reduced expression of *DAPK1*, cause an inherited susceptibility to CLL [41,42]. *DAPK1* expression of the CLL allele may be downregulated due to increased HOXB7 binding, and promotor methylation results in loss of *DAPK1* expression.

Somatic inactivation of the ataxia telangiectasia gene (*ATM*) occurs in >20% of CLL cases, and in some cases a germline mutation is also present. However, while heterozygous *ATM* mutations may confer an increased risk of CLL, germline *ATM* mutations do not make a major contribution to familial CLL [35]. Wiley et al. [43] reported that a loss-of-function polymorphism in the cytolytic P2X7 receptor gene was overrepresented in patients with CLL compared to controls. However, in another study the P2X7 SNP was associated with survival but not risk of CLL [44].

Familial clusters of hairy cell leukemia, an uncommon subtype of CLL with a prevalence of 1 per 150,000, have been reported in over 30 cases [45,46]. Linkage to specific human leukocyte antigen (HLA) haplotype has been suggested, but the influence of genetic and environmental factors in familial cases is unclear.

Specific genetic disorders associated with acute lymphoblastic leukemia include DS, NF1, chromosome breakage syndromes (ataxia telangiectasia, Bloom syndrome), Li–Fraumeni syndrome, Fanconi anemia, NF1, Kostmann syndrome, DNA repair defects, and immune deficiency disorders (severe combined, Bruton agammaglobulinemia, adenosine deaminase deficiency).

An autosomal-dominant syndrome of cerebellar ataxia, hypoplastic anemia, and predisposition to AML (associated with monosomy 7 in bone marrow cells) has been described [47].

Subtypes of leukemia which may occur commonly in certain genetic conditions include juvenile myelomonocytic leukemia (JMML), where about 10% of cases arise in children with NF1 and NS, a condition due to germline gain-of-function mutations in one of the oncogenes in the RAS signaling pathway, including *PTPN11*, *SOS1*, *RAF1*, *KRAS*, and *CBL* [48,49]. Children with NS are at increased risk of developing JMML or a myeloproliferative disorder associated with NS resembling JMML in the first weeks of life. JMML constitutes about 30% of childhood cases of myelodysplastic syndrome and 2% of leukemias [50]. Somatic mutations in *CBL*, an E3 ubiquitin protein ligase responsible for the inactivation of protein tyrosine kinases by tagging them for degradation, as with *PTPN11*, *KRAS*, *NRAS*, are well characterized in a variety of leukemias, particularly JMML and CML. Germline mutations in the *CBL* gene have been found in a very small proportion of NS patients with a predisposition to JMML [51]. Legius syndrome, due to germline mutations in the *SPRED1* gene, is characterized by café-au-lait macules, axillary and inguinal freckling and neurofibromas and schwannomas. There is an increased risk of JMML in this condition also.

In DS there is a 500-fold increased risk of acute megakaryoblastic leukemia, with an overall risk of developing leukemia of about 2%, and about 10% of DS patients have transient myeloproliferative disorder (TMD) at birth. TMD is the clonal proliferation of myeloid blasts, usually with megakaryoblastic features and presents almost exclusively in DS patients [52,53]. The symptoms vary from asymptomatic leukocytosis to severe disease causing multiorgan failure. The disorder regresses in about 80% cases, but about 20% of cases will develop AML, usually by 5 years [54]. TMD is a preleukemic state, and additional mutations of genes regulating the proliferation of the



megakaryocyte progenitors such as *GATA-1* and *RUNX-1* are needed for disease progression [55]. *GATA-1* mutations are almost universally found in the AML of children with DS, and these may arise very early in the child's life.

Children with DS are also predisposed to ALL, but much less so than to acute megakaryocytic leukemia. About 20%–33% of DS ALL cases are found to have a point mutation in exon 14 of *JAK2*, compared to 0 out of 41 *JAK2* mutations in non-DS ALL patients [56]. Patients with a *JAK2* mutation usually present earlier and with a higher WBC count than other ALL patients [57]. *JAK2* mutations confer a cell advantage by promoting cell growth via the JAK/STAT signaling pathway [58], and the mutations found in DS patients appear to be specific for DS ALL [54,59].

In non-DS patients the most common genetic association seen in ALL is the *TEL*–*AML1* translocation between chromosomes 12 and 21, which accounts for about 25% of childhood ALL [60]. The mechanism by which the fusion protein formed by this translocation causes leukemogenesis is unclear, but the roles of both *TEL* and *AML1* have been shown to be vital in hematopoiesis. The translocation disrupts parts of *AML1* known as the core-binding factor, causing disruption of normal differentiation of B-cell progenitors [61]. The translocation can be found in the blood cells of patients at birth, which is many years before the age at which ALL patients typically present, as it takes so many years for the disease to develop sufficiently to present clinically [60].

Myelofibrosis is a precursor of myeloid malignancy, and recent studies have shown that mutations in a multitude of genes can occur somatically in this condition and also probably can cause a susceptibility to this condition if a germline mutation in the gene is present. The list includes *JAK2*, *MPL*, *TET2*, *IDH1/2*, *DNMT3A*, *SH3B2* (*LNK*), *ASXL1*, *CALR*, *EZH2*, and *CBL* [62]. These genes involved can broadly be categorized into telomeropathies, DNA repair defects, biochemical defects, PAX genes, histone modifiers, methylation and splicing genes, and ribosomal disorders.

It has been suggested that leukemia susceptibility can be divided into the following main categories: genetic instability and DNA repair syndromes, RAS/MAPK pathway abnormalities, bone marrow failure syndromes, telomeropathies, immunodeficiency, transcription factor abnormalities, familial leukemia, and aneuploidy syndromes [63]. A summary table of genetic disorders associated with an increased risk of leukaemia is shown in Table 24.1.

Our understanding of inherited susceptibility to hematological malignancies is rapidly increasing and will require a wider appreciation.

---

## 24.9 Tumor markers in circulating blood

---

Increasingly clinicians are looking for noninvasive ways of diagnosing cancer. Examples of tumor markers include prostate-specific antigen for prostate cancer, cancer antigen (CA) 125 for ovarian cancer, calcitonin for medullary thyroid cancer, alpha-fetoprotein for liver cancer and human chorionic gonadotropin for germ cell tumors, such as testicular cancer and ovarian cancer. CA15-3 is a tumor marker for breast cancer. CA19-9 antigen has been shown to be elevated in the blood of some patients with gastrointestinal tumors. Carcinoembryonic antigen is useful in monitoring patients with various types of malignancies, such as gastrointestinal, pancreatic, breast, and lung cancers.

Very recently it has been reported that it is possible to detect germline DNA methylation in cancer susceptibility genes in peripheral blood. Increased methylation levels in functional gene promoters has been associated with breast cancer risk and may in the future be used as a predictive test for cancer susceptibility [24].

---

## 24.10 Genetic counseling for inherited cancer susceptibility

---

Demand for cancer risk assessment based upon the estimation of the genetic component of cancer risk to a given individual is increasing rapidly. This is both because of increased public awareness of the genetic aspects of cancer susceptibility, and as a result of requests from clinicians for evaluation of their patients' cancer risk, so that appropriate surveillance protocols can be developed. Risk prediction for common cancers is based upon careful assessment of family history of cancer and cancer-related syndromes, and a personal history and examination (where appropriate). The genetic risk assessment requires confirmation of the diagnosis in affected relatives whenever possible. Evaluation of the pedigree will allow the clinician to determine whether there is likely to be a strong inherited cancer susceptibility condition present in an affected individual in the family, allowing genetic testing to be done to try and identify the mutation causing the susceptibility in an affected individual in

the family, with downstream cascade testing offered if a mutation is detected. Alternatively, the family history may be due to a combination of less strong genetic factors and environmental influences, where genetic tests are not appropriate, but risk assessment is available for individuals in the family based on pedigree analysis, leading to appropriate surveillance and prophylactic measures being offered to individuals at increased cancer risk. Risk assessment allows the person counseled (the consultand) to be assigned to a risk category (high, moderate, or low), which then can allocate them for appropriate management. Close links with oncologists and clinicians such as in joint or multidisciplinary meetings are helpful for arranging surveillance and prophylactic measures. Education should be provided for primary care and other referral clinicians with guidelines for appropriate referrals for screening or genetic testing.

When a high-risk mutation has been identified in an individual, it is important to consider their own clinical risk-reducing management, and also who else in the family may be at risk and could benefit from genetic testing. This requires arranging for such relatives to be informed of this in an appropriate manner. In some cases the treatment of cancer may also be altered by knowing that the individual had a germline mutation in a cancer predisposing gene.

Many inherited conditions that cause an increased risk of certain specific cancers have a clear clinical phenotype, such as FAP, characterized by multiple colonic adenomas from the teenage years, and VHL disease (characterized by cerebellar tumors, renal cell cancers, and pheochromocytomas). Many of these are rare, but when diagnosed clinically there is a clear inheritance pattern and known cancer risks, with well-established screening and prophylactic protocols to reduce the associated cancer risks. Genetic counseling for such conditions involves making the correct diagnosis, identifying the causative mutation if possible, discussing risk-reducing options, and arranging predictive testing for at-risk relatives of the index case. However, in many situations there is no clear clinical phenotype to help make the diagnosis, such as in breast cancer susceptibility and Lynch syndrome, so often pedigree analysis is necessary to help make the diagnosis. The family history should be ascertained carefully, obtaining details about first-, second-, and third-degree relatives on both sides of the consultand's family, determining the age at diagnosis of cancer, histology if known, and number and types of cancer and other lesions present in each individual. If a genetic test result for a mutation in a cancer predisposing gene is available, consent should be sought from the affected individual to obtain details of this. Pathological details of the cancers in affected relatives should be confirmed if possible, as the consultand may have erroneous knowledge of their relatives' cancers. Where possible, a family history questionnaire should be sent to the consultand before their appointment, so that these details can be obtained at leisure. It is also important to obtain a complete past and current history of any illnesses in the consultand, which could indicate a cancer susceptibility, for example, freckly pigmentation of the lips in PJS, or multiple colorectal polyps in FAP. Features in the family history that indicate an increased risk for cancer in the consultand are several close relatives affected with the same cancer or a related cancer (e.g., breast and ovarian cancer) on the same side of the family, cancers diagnosed at a young age, individuals affected with more than one related cancer, or individuals with features of cancer predisposing conditions such as FAP.

Genetic counseling is the process by which the personal and family history is taken and assessed, and information derived from pedigree analysis and genetic testing is explained to the individual attending the genetic clinic (the consultand) to indicate their risk of developing a disease and their risk of handing on a genetic susceptibility to their offspring. Specific issues related to genetic counseling in the context of genetic testing are discussed in the following section.

---

## 24.11 Diagnostic and predictive genetic testing for cancer

---

Genetic testing can be divided as follows: (1) the diagnosis of a genetic disorder or disease susceptibility in the consultand by pedigree analysis, clinical examination, and genetic testing, and (2) predictive testing in an unaffected relative of an index case, by a genetic test for a known genetic alteration identified in the affected relative (the proband).

Genetic testing looks for inherited mutations or chromosomal abnormalities. Such tests may be divided into two main groups: the first includes tests for disorders that will inevitably develop, with known age-related risks of developing cancer, and where there are clearly established prophylactic and surveillance measures that can be put in place to lower the risk of cancer, such as FAP or PJS, and the second includes tests for genes conferring an inherited susceptibility to certain cancers without a clear clinical phenotype. These include Lynch syndrome (hereditary nonpolyposis colorectal cancer), and hereditary breast/ovarian cancer susceptibility due to inherited

mutations in *BRCA1* or *BRCA2*. In these conditions, it is not inevitable that the affected individual will develop cancer, but the probability that they will do so is significantly increased. There is good evidence that environmental factors influence whether cancer will develop in such individuals, and there are many lifestyle, surveillance, and prophylactic options, which can be offered to them to reduce their cancer risks, such as taking regular aspirin to reduce the risk of colorectal cancer in Lynch syndrome.

In many situations the family history may indicate that it is unlikely that a strong genetic susceptibility condition is present in the family, so the counseling session would be based on giving the consultand an estimate of their own empirical risk of developing the cancers seen in relatives, which may be increased due to shared genetic and environmental factors in the family. The risk estimate can be divided into low, moderate, and high risk, and surveillance protocols tailored to the risk estimate. The consultand should be advised to inform the genetics center if there are any changes to the family history going forward. The estimate of risk is made based on the number of affected individuals with the same or related cancer (e.g., breast and ovarian cancer), the ages at diagnosis, and the relationship to the consultand. Certain characteristics of the cancers are important, for instance, breast cancers with a “triple-negative” phenotype (estrogen and progesterone receptor negative and Her2-neu negative) are more likely to occur in women with germline *BRCA1* mutations, so a woman with this type of cancer is more likely to carry a *BRCA1* mutation, and colorectal cancers, which are microsatellite unstable and proximally sited, are commoner in Lynch syndrome than in sporadic cancers. Information may be available regarding the staining pattern in tumors using antibodies to the proteins produced by Lynch syndrome that can also help determine whether the affected individual has Lynch syndrome.

Surveillance protocols are available in many countries, which vary depending on the estimated degree of risk, and such interventions should be discussed and put in place if appropriate (e.g., mammograms annually from 40 years age, or a single colonoscopy at 55 years, depending on local guidelines). These protocols are developed based on the age-related risk of cancer and the economic cost–benefit considerations for the health service. Details of surveillance protocols of this type can be found at NICE (the United Kingdom) for breast cancer, and on the GUT website for colorectal cancer.

In many health-care systems, there are thresholds for the risk estimates above which genetic testing for a cancer susceptibility is available in the health service ([www.nice.org.uk/pdf/CG014Fullguideline.pdf](http://www.nice.org.uk/pdf/CG014Fullguideline.pdf); <http://www.cancer.gov/cancerinfo/pdq/genetics/risk-assessment-and-counseling>).

To identify a pathogenic germline mutation in a family, it is usual to start by testing a blood (or tissue) sample from an affected relative, following informed consent for testing for a genetic cancer susceptibility. This requires an initial approach from the individual being counseled, and some family and confidentiality problems can arise over this. It is essential that the affected relative understands the nature of the tests being performed and the possible emotional impact of a positive (or a negative) mutation result, and its relevance in terms of insurance and employment. When arranging diagnostic genetic tests, it is important to have a rapport with the tested individual, and to explain possible test outcomes before the test is undertaken, with a clear plan for the communication of results.

There may be specific implications of the test result for the affected individual tested. Thus an affected woman found to carry a pathogenic *BRCA1* or *BRCA2* mutation might opt for prophylactic mastectomy, more radical breast surgery at initial treatment for cancer, or prophylactic oophorectomy, or subtotal colectomy, and hysterectomy in postmenopausal women might be recommended for the treatment of early colorectal cancer in individuals with Lynch syndrome. Cancers in patients who possess germline *BRCA1/2* mutations may be more susceptible to PARP inhibitors, as the tumors are unable to repair double-strand breaks in DNA in the tumor due to the biallelic *BRCA1/2* mutations present in the tumor. For this reason, testing affected individuals is sometimes performed for management indications, with little pretest counseling. Tumors with microsatellite instability, as in individuals with Lynch syndrome (caused by germline mutations in genes encoding DNA MMR enzymes) have greater resistance to cisplatin and 5-fluorouracil than MMR-proficient patients. Early indications are that the emotional impact of being informed that they carry a susceptibility mutation with only minimal genetic counseling has little adverse effect on individuals already affected with cancer. However, some mutation carriers may experience severe guilt feelings because they may have handed on the mutation to their children who may have developed cancer, so pretest counseling should include a careful exploration of the reasons for having the test, the cancer risks in individuals carrying mutations in the gene tested, and the anticipated impact of a positive, negative, and uncertain result. The surveillance, prophylactic, and treatment options available to mutation carriers should be clearly explained, and a strategy for management in the event of a positive result clearly outlined prior to testing. It is also helpful to discuss the way in which the results are to be conveyed, to whom, and any confidentiality constraints. Where a positive result was not anticipated, such a result can be

unexpected and upsetting in an affected individual, for instance, in a young woman affected with breast cancer but with no family history of cancer, or in a woman of Ashkenazi Jewish (AJ) descent with little family history of breast/ovarian cancer, tested for a founder mutation in the *BRCA* genes (common in the AJ population). Increasingly, as tests become cheaper and more widely available, there is a case to be made for testing all individuals newly diagnosed with cancer in certain categories, such as ovarian cancer, for instance, or individuals diagnosed at a young age especially if they have features of hereditary cancers such as triple-negative breast cancer. This could influence the type of surgery or treatment offered. Here it is important to have a rapid turnaround for the test result; the timescale for testing should always be explained prior to testing.

There may be one of three outcomes of a diagnostic genetic test:

1. The test may reveal a pathogenic mutation, which explains the disease in the proband and allows predictive genetic tests to be offered to their close relatives.
2. The test may not reveal a pathogenic mutation. Where no mutation is detected, no genetic test will be available for close relatives, and no explanation will have been found for the cancer in that individual. However, it does not rule out the possibility that other (probably lower penetrance) cancer predisposing genes may have contributed to the etiology of the cancer.
3. A sequence change (or variant) may be detected in the gene tested, the significance of which may not be clear, necessitating further tests (e.g., segregation of the mutation with disease in the family, loss of the normal allele in tumor tissue, in silico and functional analysis of the variant, consultation with databases of mutations detected in other patients) to clarify this ([www.genetests.org](http://www.genetests.org)). Such variants are not uncommon. There is now a good deal of information available about such polymorphisms on websites, gleaned from the experience of other centers worldwide, and databases are available, utilizing information accrued worldwide of the pathogenicity of such gene variants, which are helpful. In such cases, it is important that the tested individual understands their results and their implications. The development of improved methods of mutation testing using sequencing has resulted in the detection of increasing numbers of such polymorphisms, and the uncertain nature of their implications is sometimes very difficult to explain. For this reason, it is helpful to mention the possibility of detecting such a variant before the test is initiated. Clearly, when the pathogenicity of a variant is unknown, it cannot be used for predictive testing in the unaffected relatives in that family.

Where an unaffected individual presents for genetic risk assessment with a family history of cancer which suggests that there may be a highly penetrant cancer susceptibility gene mutation in the family, it is generally considered inappropriate to test the unaffected individual without knowing whether there is a detectable pathogenic germline mutation in the family. In order to identify the mutation in a family, it is necessary to obtain blood (or tissue) from an affected relative, with informed consent for testing for an inherited cancer susceptibility. When a pathogenic mutation is detected, the affected individual needs to agree to the release of their results to the family to enable predictive testing to be offered to at-risk individuals in that family. Occasionally, difficulties are encountered with this, and the ethical dilemmas involved in deciding whether to release genetic test information to at-risk relatives without the consent of the individual tested (thus breaking confidentiality) are complex. This may be resolved by further discussions with the family, and considering whether the interests of the individual or the family take priority. Classically, genetic counseling for a predictive test in an unaffected individual at risk of inheriting a known mutation involves a pretest session, during which the disorder and its inheritance is fully discussed, with a detailed explanation of the likelihood that the consultand will have inherited the mutation, the cancer risks associated with the disease causing mutation, and the screening and prophylactic options available to mutation carriers to reduce their risks. Issues, such as implications for insurance cover, employment, and childbearing, should be discussed. The counselor should also explore how the consultand may react psychologically to either test result, their reasons for wishing to be tested, what they would do if found to have inherited the mutation, and who they would inform of their test results. At the time the blood is drawn, the consultand will be told when the results would be expected, and how they will be given. The results may be given at a face-to-face meeting or by telephone or letter. It is recommended that the consultand brings a confidant with them with whom they can discuss the implications of the test after the counseling session.

Problems may arise in the context of genetic testing. Sometimes certain individuals prefer not to release the results of their genetic test to their relatives, which leaves the counselor with the dilemma about whether it is right to breach their confidentiality as a "duty to warn" the at-risk relatives. Genetic tests in at-risk relatives may test an intervening relative inadvertently (since they would be an "obligate carrier"), and this requires careful handling. Thus in some cases, where the consultand is not the first-degree relative of the affected individual in the family, a positive test result in the consultand may indicate that the intervening relative (e.g., the mother of



the consultand, whose own mother is a mutation carrier) is also a mutation carrier, and it is very important that this possibility is discussed prior to testing. There should be a clear decision about how the intervening relative is to be informed about the result, preferably to include genetic counseling of that individual at the same time as the consultand. In some cases the consultand may wish to undergo testing and not inform the intervening relative, possibly because they are elderly or infirm, but sometimes because of poor family relationships. This is often a difficult counseling situation.

Genetic testing for cancer susceptibility is a paradigm for testing for susceptibility to common diseases. Such testing will become increasingly important as we understand more about the complex interplay of rare high-penetrance gene mutations and less common variants of low penetrance, and environmental and lifestyle factors in the causation of cancer. The development of strategies for genetic counseling for such diseases should take account of the importance of environmental risk factors and avoid the pitfall of genetic determinism. The provision of balanced information about the predictive value of different tests is extremely important, both for the wide variety of health-care professionals who will be required to interpret test results and for the individuals taking the tests.

## References

- [1] Giardiello FM, Trimpath JD. Peutz–Jeghers syndrome and management recommendations. *Clin Gastroenterol Hepatol* 2006;4:408–15.
- [2] Sayed MG, et al. Germline SMAD4 or BMPR1A mutations and phenotype of juvenile polyposis. *Ann Surg Oncol* 2002;9:901–6.
- [3] Zhou XP, Woodford-Richens K, Lehtonen R, et al. Germline mutations in *BMPR1A/ALK3* cause a subset of juvenile polyposis syndrome and of Cowden and Bannayan–Riley–Ruvalcaba syndromes. *Am J Hum Genet* 2001;69:704–11.
- [4] Jaiswal AS, Balusu R, Narayan S. Involvement of adenomatous polyposis coli in colorectal tumorigenesis. *Front Biosci* 2005;10:1118–34.
- [5] Ngeow J, Sesock K, Eng C. Breast cancer risk and clinical implications for germline PTEN mutation carriers. *Breast Cancer Res Treat* 2017;165:1–8.
- [6] Ghossaini M, Pharoah PDP, Easton DF. Inherited genetic susceptibility to breast cancer: the beginning of the end or the end of the beginning? *Am J Pathol* 2013;183(4):1038–51.
- [7] Goyal G, Fan T, Silberstein PT. Hereditary cancer syndromes: utilising DNA repair deficiency as therapeutic target. *Fam Cancer* 2016;15:359–66.
- [8] Lubbe SJ, Webb EL, Chandler IP, Houlston RS. Implications of familial colorectal cancer risk profiles and microsatellite instability status. *J Clin Oncol* 2009;27(13):2238–44.
- [9] Lucci-Cordisco E, Zito I, Gensini F, Genuardi M. Hereditary nonpolyposis colorectal cancer and related conditions. *Am J Med Genet* 2003;122A(4):325–34.
- [10] Peltomäki. Update on Lynch syndrome genomics. *Fam Cancer* 2016;15:385–93.
- [11] Vasen HFA. Review article: the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *Aliment Pharmacol Ther* 2007;26(Suppl. 2):113–26.
- [12] Grover S, Kastrinos F, Steyerberg EW, Cook EF, Dewanwala A, Burbidge LA, et al. Prevalence and phenotypes of *APC* and *MUTYH* mutations in patients with multiple colorectal adenomas. *J Am Med Assoc* 2012;308(5):485–92.
- [13] Short E, Thomas LE, Hurley J, et al. Inherited predisposition to colorectal cancer: towards a more complete picture. *J Med Genet* 2015. Available from: <https://doi.org/10.1136/jmedgenet-2015-103298>.
- [14] Valle L. Recent discoveries in the genetics of familial colorectal cancer and polyposis. *Clin Gastroenterol Hepatol* 2017;15(6):809–19.
- [15] Guillotin D, Martin SA. Exploiting DNA mismatch repair deficiency as a therapeutic strategy. *Exp Cell Res* 2014;329:110–15.
- [16] Wimmer K, Beilken A, Nustede R, et al. A novel germline POLE mutation causes an early onset cancer prone syndrome mimicking constitutional mismatch repair deficiency. *Fam Cancer* 2017;16:67–71.
- [17] Zhang M, Liu Y, Fu J, Hu Y, Sun Z. Functional analysis of RET with multiple endocrine neoplasia type 2. *J Cancer Res Ther* 2017;13(5):823–8.
- [18] Villani A, Greer MC, Kalish JM, Nakagawara A, Nathanson KL, Pajtlar KW, et al. Recommendations for cancer surveillance in individuals with RASopathies and other rare genetic conditions with increased cancer risk. *Clin Cancer Res* 2017;23(12):e83–90. Available from: <https://doi.org/10.1158/1078-0432.CCR-17-0631>.
- [19] Gossage L, Eisen T, Maher ER. VHL, the story of a tumour suppressor gene. *Nat Rev Cancer* 2015;15:55–64.
- [20] Welander J, Andreasson A, Juhlin CC, Wiseman RW, Bäckdahl M, Höög A, et al. Rare germline mutations identified by targeted next-generation sequencing of susceptibility genes in pheochromocytoma and paraganglioma. *J Clin Endocrinol Metab* 2014;99(7):E1352–60. Available from: <https://doi.org/10.1210/jc.2013-4375>.
- [21] Klein S, Lee H, Gahremani S, et al. Expanding the phenotype of mutations in *DICER1*: mosaic missense mutations in the RNase IIIb domain of *DICER1* cause GLOW syndrome. *J Med Genet* 2015;51:294–302.
- [22] Park KS, Mitra A, Rahat B, Kim K, Pfeifer K. Loss of imprinting mutations define both distinct and overlapping roles for misexpression of IGF2 and of H19 lncRNA. *Nucleic Acids Res* 2017;45(22):12766–79. Available from: <https://doi.org/10.1093/nar/gkx896>.
- [23] Boldry EJ, Patel YM, Kotapati S, Esades A, Park SL, Tiirikainen M, et al. Genetic determinants of 1,3-butadiene metabolism and detoxification in three populations of smokers with different risks of lung cancer. *Cancer Epidemiol Biomarkers Prev* 2017;26(7):1034–42.
- [24] Joo JE, Dowty JG, Kilne RL, et al. Heritable DNA methylation marks associated with susceptibility to breast cancer. *Nat Commun* 2018;9(867):1–12.
- [25] Tung N, Domchek SM, Stadler Z, et al. Counselling framework for moderate-penetrance cancer-susceptibility mutations. *Nat Rev Clin Oncol* 2016;13:581–8.



- [26] Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med* 2008;358(26):2796–803. Available from: <https://doi.org/10.1056/NEJMsa0708739>.
- [27] Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. *Nat Genet* 2008;40(1):17–22.
- [28] Kwong YL, Ng MH, Ma SK. Familial acute myeloid leukemia with monosomy 7: late onset and involvement of a multipotential progenitor cell. *Cancer Genet Cytogenet* 2000;116:170–3.
- [29] Ganly P, Walker LC, Morris CM. Familial mutations of the transcription factor RUNX1 (AML1, CBFA2) predispose to acute myeloid leukaemia. *Leuk Lymphoma* 2004;45(1):1–10.
- [30] Ernst T, Chase AJ, Score J, et al. Activating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet* 2011;42:722–5.
- [31] Park S, Picard F, Dreyfus F. Erythroleukemia: a need for a new definition. *Leukemia* 2002;16:1399–401.
- [32] Kurzrock R, Guterman JU, Talpaz M. The molecular genetics of Philadelphia chromosome-positive leukaemias. *N Engl J Med* 1988;319:990–8.
- [33] Laurent EL, Talpaz M, Kantarjian H, Kurzrock R. The BCR gene and Philadelphia chromosome-positive leukaemogenesis. *Cancer Res* 2001;61:2343–55.
- [34] Goldin LR, Slager SL, Caporaso NE. Familial chronic lymphocytic leukaemia. *Curr Opin Haematol* 2010;17:350–5.
- [35] Houlston RS, Sellick G, Yuille M, Matutes E, Catovsky D. Causation of chronic lymphocytic leukemia – insights from familial disease. *Leuk Res* 2003;27:871–6.
- [36] Rawstron AC, Yuille MR, Fuller J, Cullen M, Kennedy B, Richards SJ, et al. Inherited predisposition to CLL is detectable as subclinical monoclonal B-lymphocyte expansion. *Blood* 2002;100(7):2289–90.
- [37] Slager S, Caporaso NE, Silvia de Sanjose S, Goldin L. Genetic susceptibility to chronic lymphocytic leukemia. *Semin Hematol* 2013;50(4). Available from: <https://doi.org/10.1053/j.seminhematol.2013.09.007>.
- [38] Crowther-Swanepoel D, Broderick P, Di Bernardo MC, Dobbins SE, Torres M, Mansouri M, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukaemia risk. *Nat Genet* 2010;42:132138.
- [39] Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet* 2008;40(10):1204–10. Available from: <https://doi.org/10.1038/ng.219> Epub 2008 Aug 31.
- [40] Houlston RS. Low penetrance susceptibility to haematological malignancy. *Curr Opin Genet Dev* 2010;20:245–50.
- [41] Lynch HT, Ferrara KM, Weisenburger DD, Sanger WG, Lynch JF, Thomé SD. Genetic counseling for DAPK1 mutation in a chronic lymphocytic leukemia family. *Cancer Genet Cytogenet* 2008;186(2):95–102.
- [42] Raval A, Tanner SM, Byrd JC, Angerman EB, Perko JD, Chen SS, et al. Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. *Cell* 2007;129(5):879–90.
- [43] Wiley JS, et al. A loss-of-function polymorphic mutation in the cytolytic P2X7 receptor gene and chronic lymphocytic leukaemia: a molecular study. *Lancet* 2002;359(9312):1114–19.
- [44] Thunberg U, Tobin G, Johnson A, Söderberg O, Padyukov L, Hultdin M, et al. Polymorphism in the P2X7 receptor gene and survival in chronic lymphocytic leukaemia. *Lancet* 2002;360(9349):1935–9.
- [45] Colovic MD, Jankovic GM, Wiernik PH. Hairy cell leukemia in first cousins and review of the literature. *Eur J Haematol* 2001;67(3):185–8.
- [46] Cetiner M, Adiguzel C, Argon D, Ratip S, Eksioğlu-Demiralp E, Tecimer T, et al. Hairy cell leukemia in father and son. *Med Oncol* 2003;20(4):375–8.
- [47] Daghistani D, Toledano SR, Curless R. Monosomy 7 syndrome. *Cancer Genet Cytogenet* 1990;44:263–9.
- [48] La Starza R, Rosati R, Roti G, Gorello P, Bardi A, Crescenzi B, et al. A new NDE1/PDGFRB fusion transcript underlying chronic myelomonocytic leukaemia in Noonan syndrome. *Leukemia* 2007;21(4):830–3 Epub 2007 Feb 15.
- [49] Pandit B, Sarkozy A, Pennacchio LA, Carta C, Oishi K, Martinelli S, et al. Gain-of-function RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. *Nat Genet* 2007;39(8):1007–12 Epub 2007 Jul 1.
- [50] Jongmans MC, van der Burgt I, Hoogerbrugge PM et al., Cancer risk in patients with Noonan syndrome carrying a PTPN11 mutation, *Eur J Hum Genet* 2011;19(8):870–874.
- [51] Niemeyer CM, Kang MW, Shin DH, Furlan I, Erlacher M, Bunin NJ, et al. Germline CBL mutations cause developmental abnormalities and predispose to juvenile myelomonocytic leukaemia. *Nat Genet* 2010;42:794–800.
- [52] Izraeli S, Rainis L, Hertzberg L, Smooha G, Birger Y. Trisomy of chromosome 21 in leukemogenesis. *Blood Cells Mol Dis* 2007;39(2):156–9.
- [53] Klusmann JH, Creutzig U, Zimmerman M, Dworzak M, Jorch N, Langebrake C, et al. Treatment and prognostic impact of transient leukaemia in neonates with Down syndrome. *Blood* 2007;111(6):2991–8.
- [54] Rabin KR, Whitlock JA. Malignancy in children with trisomy 21. *Oncologist* 2009;14(2):164–73.
- [55] Izraeli S. Perspective: chromosomal aneuploidy in leukaemia – lessons from Down syndrome. *Haematol Oncol* 2005;24(1):3–6.
- [56] Kearney L, De Castro DG, Yeung J, Procter J, Horsley SW, Eguchi-Ishimae M, et al. Specific JAK2 mutation (JAK2R683) and multiple gene deletions in Down syndrome acute lymphoblastic leukaemia. *Blood* 2009;113(3):646–8.
- [57] Tigay JH. A comparison of acute lymphoblastic leukaemia in Down syndrome and non-Down syndrome children: the role of trisomy 21. *J Paediatr Nurs* 2009;26(6):362–8.
- [58] Mulligan CG. JAK2 – a new player in acute lymphoblastic leukaemia. *Lancet* 2008;372(9648):1448–50.
- [59] Bercovich D, Ganmore I, Scott LM, Wainreb G, Birger Y, Elimelech A, et al. Mutations of JAK2 in acute lymphoblastic leukaemias associated with Down syndrome. *Lancet* 2008;372(9648):1484–92.
- [60] Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukaemia. *J Clin Oncol* 2005;23(26):6306–15.
- [61] Ford AM, Palmi C, Bueno C, Hong D, Cardus P, Knight D, et al. The TEL-AML1 leukaemia fusion gene dysregulates the TGF- $\beta$  pathway in early B lineage progenitor cells. *J Clin Invest* 2009;119(4):826–36.

- [62] Tabarrok A, Tiu RV. Molecular genetics of myelofibrosis and its associated disease phenotypes. *Transl Med* 2014;8(7):53–64 ISSN 2239-9747 2014.
- [63] Saida S. Predispositions to leukemia in Down syndrome and other hereditary disorders. *Curr Treat Options Oncol* 2017;18:41.

## Further reading

- Altieri A, Bermejo JL, Hemminki K. Familial risk for non-Hodgkin lymphoma and other lymphoproliferative malignancies by histopathologic subtype: the Swedish Family-Cancer Database. *Blood* 2005;106:668–72 [PubMed: 15811955].
- Chakravarti A, Halloran SL, Bale SJ, Tucker MA. Etiological heterogeneity in Hodgkin's disease: HLA linked and unlinked determinants of susceptibility independent of histological concordance. *Genet Epidemiol* 1986;3:407–15.
- Chang ET, Smedby KE, Hjalgrim H, et al. Family history of haematopoietic malignancy and risk of lymphoma. *JNCI* 2005;97:1466–74.
- Comotti B, Bassan R, Buzzeti M, Finazzi G, Barbui T. Multiple myeloma in a pair of twins. *Br J Haematol* 1987;65:123–4.
- Davies HD, Leusink GL, McConnell A, Deyell M, Cassidy SB, Fick GH, et al. Myeloid leukemia in Prader–Willi syndrome. *J Pediatr* 2003;142:174–8.
- Ding J, Komatsu H, Wakita A, et al. Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. *Blood* 2004;103:4198–200.
- Goldin LR, Pfeiffer RM, Hemminki K. Familial risk of lymphoproliferative tumours in families of patients with chronic lymphocytic leukaemia: results from the Swedish Family-Cancer Database. *Blood* 2004;104(6):1850–4.
- Horwitz M, Benson KF, Li FQ, Wolff J, Leppert MF, Hobson L, et al. Genetic heterogeneity in familial acute myelogenous leukemia: evidence for a second locus at chromosome 16q21–23.2. *Am J Hum Genet* 1997;61:873–81.
- Ishibe N, Sgambati MT, Fontaine L, et al. Clinical characteristics of familial B-CLL in the National Cancer Institute Familial Registry. *Leuk Lymphoma* 2001;42:99–108.
- Lynch HT, Weisenburger DD, Quinn-Laquer B, Watson P, Lynch JF, Sanger WG. Hereditary chronic lymphocytic leukemia: an extended family study and literature review. *Am J Med Genet* 2002;115:113–17.
- Menko FH, Kaspers GL, Meijer GA, Claes K, van Hagen JM, Gille JJ. A homozygous MSH6 mutation in a child with café-au-lait spots, oligodendroglioma and rectal cancer. *Fam Cancer* 2004;3(2):123–7.
- Powell SM, Zilz N, Beazer-Barclay Y, et al. APC mutations occur early during colorectal tumorigenesis. *Nature* 1992;359:235–7.
- Smith ML, Cavenagh JD, Lister TA, Fitzgibbon J. Mutation of CEBPA in familial acute myeloid leukemia. *N Engl J Med* 2004;351:2403–7.
- Wiemels JL, de Smith AJ, Xiao J, Lee ST, Muench MO, Fomin ME, et al. A functional polymorphism in the CEBPE gene promoter influences acute lymphoblastic leukemia risk through interaction with the hematopoietic transcription factor Ikaros. *Leukemia* 2016;30(5):1194–7. Available from: <https://doi.org/10.1038/leu.2015.251>.

# Clinical molecular nephrology—acute kidney injury and chronic kidney disease

*Shivaram Hegde*

Children Kidney Centre, University Hospital of Wales, Cardiff, United Kingdom

## 25.1 Introduction

Recent developments in genomic understanding and technology have advanced our knowledge about molecular pathogenesis of a variety of renal disorders. Genetic influences might modify the development of Acute kidney Injury (AKI) and are also thought to affect the severity and outcome of AKI. Inherited kidney diseases are one of the common cause of Chronic Kidney Disease (CKD), particularly in children. Recent advances have witnessed identification genes affecting the susceptibility to CKD.

Inherited kidney diseases are heterogeneous and relatively rare poorly defined group of disorders. They are responsible for approximately 10% of adult end-stage renal disease (ESRD) [1] and at least 70% of chronic kidney disease (CKD) in children [2]. Genetic testing is currently recommended for patients with early-onset nephropathy and/or other clinical features consistent with an inherited form of disease as well as for evaluation of living kidney donors [3]. The identification of a disease causing mutation in patients with CKD has several clinical implications, including the ability to provide a definitive diagnosis, potential to place the clinical phenotype into context by gene-specific stratification and through the delivery of personalized medicine, the ability to provide precise genetic counseling, to diagnose unrecognized affected family members, to avoid unnecessary diagnostic procedures and treatments, to detect and treat asymptomatic (or subtle) extrarenal manifestations early, to provide guidance for monitoring of potential future complications, and to guide advanced medical management on a gene-specific basis [2]. Common modalities for diagnostic genetic testing include Sanger sequencing, chromosomal microarray, targeted next-generation sequencing (NGS) panels, whole-exome sequencing (WES), and whole-genome sequencing (WGS) [3].

## 25.2 Acute kidney injury

Acute kidney injury (AKI) is common in hospitalized patients particularly those who are critically ill and is associated with prolonged hospital stay and increased mortality. The long-term risk of developing CKD and ESRD after AKI is high, even after initial recovery of renal function. Currently, there is no proven and effective therapy, and the known clinical risk factors alone do not fully explain the risk of developing AKI.

Patients with similar injury to the kidneys can progress differently with different outcomes, which suggest that genetic risk factors might influence the development of AKI. Genetic differences may modulate the renal response to an insult, determining whether an individual progress to more severe injury or recovery [4]. Identifying genetic risk factors for AKI could help to develop personalized AKI-specific therapy.

AKI is measured by increasing serum creatinine levels and decreasing urine output. Serum creatinine however is not an ideal biomarker as the changes in its serum levels are not specific to a specific pathophysiology or the anatomical structure affected by the injury. Creatinine level does not rise until substantial kidney function has

been lost, which might take several hours or days, hence always lagging structural changes that occur in the kidney during the early stage of AKI.

Recent advances in genomic, metabolomic, and proteomic research have resulted in development of novel *biomarkers* of renal function. Biomarkers can be used clinically to screen for, diagnose, or monitor the activity of diseases and to guide molecularly targeted therapy or assess therapeutic response [5].

Various protein biomarkers, such as neutrophil gelatinase-associated lipocalin, interleukin-18 (IL-18), kidney injury molecule 1, liver-type fatty acid-binding protein, insulin-like growth factor-binding protein 7  $\times$  tissue inhibitor of metalloproteinases-2, calprotectin, and urinary angiotensinogen, have been described during the last decade [6]. Recent advances in the field of omics technologies are likely to add other types of biomolecules, such as RNAs, lipids, and metabolites as biomarkers. However, most of these biomarkers are poorly defined by their AKI-associated molecular context [7]. Although biomarkers could help to understand the pathophysiological mechanisms and some allowed early diagnosis in some forms of AKI, their clinical utility and relevance in AKI in general is unclear.

### 25.2.1 Candidate gene association studies

Many genes and variants already hypothesized to play a role in AKI, which have been used by most published studies addressing the genetic risk factors for AKI. Some of the genes/at risk variants studied include apolipoprotein E(alleles E2, E3, and E4) [8,9]; renin-angiotensin system (variants angiotensin-converting enzyme, angiotensin II receptor type 1 and angiotensinogen) [10,11]; proinflammatory pathways (tumor necrosis factor alpha, IL-1 beta, IL-6, IL-8, interferon gamma, transforming growth factor beta and IL-10) [9,12,13]; NOS3 (which encodes endothelial nitric oxide synthase) [14]; vascular endothelial growth factor A [15]; antimicrobial activity and reactive oxygen species: cytochrome b-245 alpha (CYBA), catalase (CAT), and myeloperoxidase (MPO) [16]; and catecholamine synthesis and metabolism (catechol-O-methyltransferase (COMT) and Phenylethanolamine N-Methyltransferase (PNMT)) [17].

These studies frequently found conflicting results and have failed to identify contributory variants consistently [4].

### 25.2.2 Genome-wide association studies

Data from a number of genome-wide association studies (GWASs) allow for an unbiased search for correlation between single-nucleotide polymorphisms (SNPs) and phenotypes, which may be better suited to the complexity of AKI [18]. GWASs use genomic variations, termed SNPs, to identify regions of the genome associated with the disease status or a clinical phenotype [19]. A GWAS identified two new loci in *BBS9* and the *GRM7*/*LMCD1-AS1* intergenic region associated with postcoronary bypass graft surgery AKI [20]. This association might provide greater insight into the pathogenesis of AKI.

Genomics have immensely helped in diagnosis and treatment of *atypical hemolytic uremic syndrome* (aHUS). AKI is the most common finding at initial presentation of aHUS and leads to CKD in many who recover from the acute episode. aHUS is a disease characterized by complement overactivation related to inherited defects in complement genes or acquired autoantibodies against complement regulatory proteins [21]. Modern genetic testing has identified the causative mutations in complement genes, which results in complement overactivation. This information led to development of targeted, complement inhibitor drug, eculizumab. Confirmed mutations in complement genes usually predict a good response to eculizumab, while only a subset of patients with non-complement gene mutation such as *DGKE* (diacylglycerol kinase epsilon) are responsive. Eculizumab has transformed the outcome of aHUS and quality of life both in adults and children.

## 25.3 Chronic kidney disease

The prevalence of CKD is growing worldwide both in adults [22] and children [23], and it is associated with decreased quality of life, poor outcomes, and higher economic costs [24]. Progression to ESRD necessitates renal replacement therapy (dialysis or transplantation). Patients with CKD often present late and earlier, and more specific diagnoses are needed to manage them efficiently.

### 25.3.1 Causes of chronic kidney disease

The heritability of glomerular filtration rate is estimated to be ~30%–60% in the general population [25], and 10%–29% of adult patients with ESRD report a positive family history [26]. The age at presentation, disease course, prognosis, principle management, response to certain therapies, and outcome post–renal transplantation varies between inherited and acquired causes. However, they might be inseparable clinically and obvious only on genetic testing.

Inherited kidney diseases are responsible for approximately 10% of adult ESRD [1] and at least 70% of CKD in children [2]. Approximately 20% of cases of CKD that manifest before the age of 25 years are caused by single-gene mutations in any one of >200 different genes [2].

Diabetes (40% of CKD cases) and hypertension (28% of cases) are the two common causes for CKD in adults [27]. The primary causes of CKD however are different in children and include congenital anomalies of the kidneys and urinary tract (CAKUT—49.1% of cases), steroid-resistant nephrotic syndrome (SRNS—10.4% of cases), chronic glomerulonephritis (8.1% of cases), and renal cystic ciliopathies (5.3% of cases) [28].

Increasing utilization of *genetic testing* in nephrology has helped to detect genetic etiologies for CKD not just in children and but also in adults. The values of genetic testing [29] in nephrology are manifold and include the following.

### 25.3.2 Establishing a molecular genetic diagnosis of chronic kidney disease

Genetic testing not only confirms the diagnosis but also aids in its proper classification and appropriate management.

#### 25.3.2.1 Clinical heterogeneity in chronic kidney disease

Diagnosis of a genetic cause enables more precise distinction between *overlapping clinical phenotypes*, which can predict prognosis and guide management [30], as exemplified by the utility of genetic testing in distinguishing Alport syndrome (AS) from the related but milder condition of thin basement membrane disease (TBMD). Although both the nephropathies cause hematuria and result from mutations in COL4A3–5, they differ greatly in their outcome. AS presents in pediatric age-group with hematuria with proteinuria, progresses to ESRD, and affects other organs (eyes and ears). TBMD is a milder nonprogressive disease characterized by minimal proteinuria and normal renal function in addition to hematuria. Genotype-level knowledge can help predict phenotype: in general, biallelic COL4A3/COL4A4 or hemizygous COL4A5 mutations result in AS, whereas heterozygous COL4A3-4 mutations result in TBMD [31]. Precise molecular diagnosis, hence, not only offers valuable prognostic information but also enables planning for appropriate management of resultant CKD and assessment of extrarenal manifestations.

#### 25.3.2.2 Genetic heterogeneity in chronic kidney disease

Establishing molecular diagnosis of *genetically heterogenous conditions*, such as nephrotic syndrome and renal ciliopathies (by NGS), has major implications for clinical management and predicting prognosis. Genetic forms of nephrotic syndrome generally do not respond to immunosuppressive therapy (including steroids) and progress relatively faster to ESRD. Monogenic mutations in over 30 genes have been identified [32] as the etiology in around a third of patients with SRNS. It does not recur in the transplanted kidney as do some of the nongenetic steroid-resistant forms of nephrotic syndrome. Genetic testing may also uncover a form of SRNS that is amenable to nonimmunosuppressive treatment such as mutations in the CoQ10 biosynthetic pathway [32].

Patients with ciliopathies have gene mutations that encode components of the ciliary apparatus many of whom progress to ESRD. NGS enables accurate diagnosis of different types of renal ciliopathies and prompts further assessment for extrarenal manifestations based on the genetic form discovered, such as visual impairment and abnormal hepatic function [33].

### 25.3.3 Understanding pathogenesis of chronic kidney disease

Molecular diagnosis of various genetic renal diseases also provides with some insight into the disease mechanisms of various renal diseases.



Modern sequencing methods (NGS) have identified causal genes for a variety of clinical phenotypes, such as renal ciliopathies, CAKUT [34], and focal segmental glomerulosclerosis (FSGS)/SRNS, which has helped to understand the disease mechanism.

The molecular pathogenesis of disease, such as membranoproliferative glomerulonephritis and aHUS, is well understood following discovery of variants in the DGKE gene, which encodes an intracellular kinase expressed in endothelium, platelets, and podocytes, among individuals with these diseases. This explains how signaling disruptions lead to endothelial damage, complement, and podocyte dysregulation [35,36].

### 25.3.4 Genetic variants and chronic kidney disease

Copy number variants (CNVs) have been found (array comparative genomic hybridization) to contribute substantially to CAKUT [37] and are also noted in pediatric CKD patients of various etiologies. Recent studies have suggested significant genetic overlap between the CNVs found in children with nephropathy and those in children with neurodevelopmental disorders [38]. Neurocognitive dysfunction has been associated with renal dysfunction in children and adults [39], and the genetic overlap noted suggests this dysfunction may in fact be part of a greater multisystem disorder. Thus genetic testing in nephrology has the potential to identify extrarenal manifestations that may otherwise be missed or dismissed as secondary complications and enable physicians to provide patients with more specialized and effective care [29].

Genetic variants can increase the risk of CKD development and progression of CKD to ESRD in certain racial groups. African-Americans develop ESRD at rates four to five times higher than European-Americans. Sequence variants in the APOL1 gene on chromosome 22 were associated with large increases in renal disease risk in African-Americans [40]. These APOL1 disease-associated variants increase the risk of hypertension-attributed ESRD and FSGS by 7–10-fold [41].

The renal epithelial cells produce uromodulin, Tamm–Horsfall protein encoded by *UMOD* gene, the most abundant protein excreted in the urine. Recent GWASs [42] have identified common SNP variants in the *UMOD* gene in association with the estimated glomerular filtration rate and the risk of developing CKD. The presence of the *UMOD* SNP rs4293393 was found to be associated with uromodulin levels, and elevated uromodulin levels preceded the development of CKD. Uromodulin overexpression led to salt-sensitive hypertension by upregulating Na–K–Cl transporter (NKCC2) phosphorylation and contributed to renal damage [43].

## 25.4 Clinical renal genomic medicine

The impact of genomic applications in the clinical molecular nephrology practice is beginning to be appreciated. In this section, few pertinent aspects are highlighted for reader to assist further reading and searching the nephrology literature.

### 25.4.1 Targeted gene panel analysis in chronic kidney disease

This involves testing patients for a set of genes that are commonly associated with the phenotype under consideration; for example, a patient with nephrotic syndrome would be tested using a panel containing genes that are commonly implicated in hereditary forms [2]. NGS is often the diagnostic genetic testing modality utilized as it provides rapid and inexpensive sequencing at higher coverage than that achieved with WES or WGS [44]. Such panels have been advocated as a first-line test for the molecular diagnosis of inherited nephropathies [2].

If the NGS targeted panel testing is negative, next option is to go for another panel with broader content or proceed directly to WES or WGS. This sequential procedure may be the most comprehensive and cost-effective approach at present, particularly, among patients whose presentation is strongly suggestive of a specific category of genetic disease [44].

Targeted gene panels are a sensitive diagnostic tool for common causes of CKD, including nephrotic syndrome [45], nephrolithiasis [46], nephronophthisis-related ciliopathies [47], and CAKUT [34]. Such targeted testing is particularly well suited to diseases that have fairly low-genetic/phenotypic heterogeneity.

Bullich et al. [48] utilized a comprehensive kidney disease gene panel (targeted NGS of 140 genes causative of or associated with cystic or glomerular nephropathies) to study 421 patients (a validation cohort of 116 patients with previously known mutations, and a diagnostic cohort of 207 patients with suspected inherited cystic disease and 98 patients with glomerular disease). In the diagnostic cohort, causative mutations were found in 78% of patients with cystic disease and 62% of patients with glomerular disease, mostly familial cases, including CNVs. They described different cystic and glomerular inherited diseases and related them to age of onset; perinatal, pediatric, or adult disease onset. Of all the genetically diagnosed patients, 15% were referred with an unspecified clinical diagnosis and in 2% genetic testing changed the clinical diagnosis. Genetic analysis was crucial to establish the correct diagnosis in 17% of cases. Etiologic diagnosis was possible in 75% of patients, the risk of the progression of renal disease and risk of extrarenal features could also be more precisely estimated. This information enables identification of at-risk relatives and provision of genetic counseling if indicated.

### 25.4.2 Targeted clinical management

Establishing a genetic diagnosis can make a significant difference to their clinical management [2].

1. Renal biopsy is not indicated in patients with confirmed genetic diagnosis of congenital or infantile nephrotic syndrome secondary to NPHS1 or NPHS2 mutations or in patients with a characteristic nephronophthisis phenotype and NPHP1 mutations.
2. Most patients with a single-gene cause of nephrotic syndrome do not respond to therapy with steroids and other immunosuppressive therapy. The adverse effects associated with the unnecessary use of these drugs are thus could be avoided.

They also do not recur in renal transplant; and hence, aggressive pretransplant antirecurrence therapies should be avoided.

### 25.4.3 Personalized pharmacotherapy in chronic kidney disease—pharmacogenomics

Currently, there are no specific therapies for most patients with CKD.

1. Salt-sensitive hypertension could result from uromodulin overexpression-related upregulated Na–K–Cl transporter activity. This suggests pharmacological inhibition of NKCC2 would be more effective in lowering blood pressure in hypertensive patients who were homozygous for UMOD promoter risk variants than in other hypertensive patients, and the uromodulin might be a therapeutic target for lowering blood pressure and preserving renal function [43].
2. Patients with SRNS due to mutation in genes involved in CoQ10 biosynthesis could be benefitted by CoQ10 supplementation [49]. CoQ10 deficiency might lead to nephrotic syndrome through the excess production and accumulation of reactive oxygen species [50]. A recent case report suggested that early initiation of the treatment, immediately after the onset of renal symptoms, was beneficial in reducing proteinuria in a patient with nephrotic syndrome caused by recessive mutations in COQ2 [51].
3. Recessive mutations in CTNS gene establish a diagnosis of cystinosis, which needs therapy with cystine-reducing drugs.

### 25.4.4 Targeted clinical surveillance

1. Patients with SRNS caused by WT1 mutations should be monitored closely as they are at increased risk of developing Wilms tumor.
2. Patients with renal ciliopathies might have associated rod-cone dystrophy and need long-term monitoring for visual impairment.

### 25.4.5 Genetic counseling

With precise genetic diagnosis in place families could be offered appropriate genetic counseling by accurately predicting disease recurrence (e.g., nephropathic cystinosis). Preimplantation genetic diagnosis might also be possible in selected cases.

## 25.5 Molecular basis of kidney transplantation

---

### 25.5.1 Human leukocyte antigen typing for renal transplantation

Human leukocyte antigen (HLA) typing of donor and recipient is essential component of transplantation as HLA plays a major role in the immune responses that determine the outcome of a transplant. HLA gene sequences are extremely polymorphic and typing, until recently it was done by a complex cellular-based assay. They are now replaced by genomic DNA testing that allows definition of a genotype to determine the HLA type at the DNA level. The introduction of NGS strategies now allows more refined, high-quality, and faster assessment of donor compatibility (HLA typing) at a lower cost.

### 25.5.2 Selection of donors

There is limited but useful information available on the genetic selection of renal transplantation donors. Monogenic forms of FSGS do not usually recur in the transplanted kidney unlike the idiopathic forms (around 30% risk). Family members of a patient with INF2 mutation will not be able to be a donor if they also have the INF2 mutations.

### 25.5.3 Predicting the long-term kidney allograft function

Genomic techniques have made some inroads for predicting the long-term outcome of renal transplantation. It is believed that cell surface antigens in addition to HLA might behave as the stimulus and target for the antiallograft immune response and affect long-term renal transplant outcome. A recent study Mesnard et al. [52] suggested that amino acid mismatches in the transmembrane proteins in individual donor/recipient pair could be a strong predictor of long-term kidney allograft function. This involved exome sequencing of DNA from both donor and recipient to know the number of amino acid mismatches in transmembrane protein [allogenomics mismatch score (AMS)], as an indicator of all possible cell surface antigen mismatches between donor and recipient. The results of the study showed that the AMS was a strong and robust predictor of long-term kidney allograft function.

## 25.6 Conclusion

---

Recent developments in genomic sequencing approaches have advanced our knowledge in nephrology by way of better characterization, new discovery, and increasing insight into molecular pathogenesis of various kidney diseases. NGS has revolutionized genomic nephrology research and drastically changed the way gene analysis is performed, enabling simultaneous sequencing of hundreds of genes. The values of genetic testing in nephrology are of many folds in clinical practice including offering precise diagnosis, insight into pathogenesis of the disease, reducing the need for invasive diagnostic procedures, ability to deliver personalized medicine, choice of treatment, disease surveillance, family counseling, and prenatal testing. Incorporating NGS in the diagnosis of nephrogenetic diseases in clinical practice will require precise disease phenotyping, functional testing using appropriate disease model and data integration.

Candidate gene association studies in AKI have found conflicting results and failed to identify contributory variants consistently. GWASs have shown some promise that might offer better insight into the pathogenesis of AKI in future. NGS has enabled positive identification of more than 200 monogenic genes as cause of early-onset CKD. Molecular diagnosis will help in etiology-based classification and prevention of CKD, in addition to the other benefits summarized earlier. Genome sequencing is of immense value in HLA typing pretransplantation and kidney donor selection (for live-related transplantation) for certain renal disorders.

Indication-driven target panel testing and drawing up evidence-based guidelines regarding clinical utility of genetic testing will be of value in avoiding some of problems inherent to genome sequencing (such as incidental findings), while improving clinical outcomes.

### Learning points

1. Inherited kidney diseases are responsible for approximately 10% of adult ESRD and at least 70% of CKD in children.
2. Recent developments in genomic sequencing approaches have advanced our knowledge about molecular pathogenesis of inherited renal disorders.
3. NGS has enabled technological progression from single-gene analysis to simultaneous sequencing of hundreds of genes.
4. The values of genetic testing in nephrology are of many folds in clinical practice including offering precise diagnosis, insight into pathogenesis of the disease, reducing the need for invasive diagnostic procedures, ability to deliver personalized medicine and choice of treatment.
5. Candidate gene association studies in AKI have failed to identify contributory variants consistently. GWASs might offer better insight into the pathogenesis of AKI in future.
6. NGS has enabled positive identification of more than 200 monogenic genes as cause of early-onset CKD. Molecular diagnosis will help in etiology-based classification and prevention of CKD, in addition to the other benefits.
7. Genome sequencing is of immense value in HLA typing pretransplantation and kidney donor selection (for live-related transplantation) for certain renal disorders.

### References

- [1] Devuyst O, Knoers NV, Remuzzi G, Schaefer F, Board of the Working Group for Inherited Kidney, Diseases of the European Renal Association, et al. Rare inherited kidney diseases: challenges, opportunities, and perspectives. *Lancet* 2014;383(9931):1844–59.
- [2] VVivante A, Hildebrandt F. Exploring the genetic basis of early-onset chronic kidney disease. *Nat Rev Nephrol* 2016;12:133–46.
- [3] Groopman EE, Rasouly HM, Gharavi AG. Genomic medicine for kidney disease. *Nat Rev Nephrol* 2018;14(2):83–104.
- [4] Laracha DB, Engoren MC, Schmidt EM, Heungd M. Genetic variants and acute kidney injury: a review of the literature. *J Crit Care* 2018;44:203–11.
- [5] Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. The case for early detection. *Nat Rev Cancer* 2003;3(4):243–52. Available from: <https://doi.org/10.1038/nrc1041>.
- [6] Kashani K, Cheungpasitporn W, Ronco C. Biomarkers of acute kidney injury: the pathway from discovery to clinical adoption. *Clin Chem Lab Med* 2017;55(8):1074–89. Available from: <https://doi.org/10.1515/cclm-2016-0973>.
- [7] Marx D, Metzger J, Pejchinovski M, Gil RB, Frantzi M, Latosinska A, et al. Proteomics and metabolomics for AKI diagnosis. *Semin Nephrol* 2018;38(1):63–87. Available from: <https://doi.org/10.1016/j.semnephrol.2017.09.007>.
- [8] Isbir SC, Tekeli A, Ergen A, Yilmaz H, Ak K, Civelek A, et al. Genetic polymorphisms contribute to acute kidney injury after coronary artery bypass grafting. *Heart Surg Forum* 2007;10(6):E439–44. Available from: <https://doi.org/10.1532/HSF98.20071117>.
- [9] Boehm J, Eichhorn S, Kornek M, Hauner K, Prinzing A, Grammer J, et al. Apolipoprotein E genotype, TNF- $\alpha$  308G/A and risk for cardiac surgery associated-acute kidney injury in Caucasians. *Ren Fail* 2014;36(2):237–43. Available from: <https://doi.org/10.3109/0886022X.2013.835267>.
- [10] McBride WT, Prasad PS, Armstrong M, Patterson C, Gilliland H, Drain A, et al. Cytokine phenotype, genotype, and renal outcomes at cardiac surgery. *Cytokine* 2013;61(1):275–84. Available from: <https://doi.org/10.1016/j.cyto.2012.10.008>.
- [11] Cardinal-Fernández P, Ferruelo A, El-Assar M, Santiago C, Gómez-Gallego F, Martín-Pellicer A, et al. Genetic predisposition to acute kidney injury induced by severe sepsis. *J Crit Care* 2013;28(4):365–70. Available from: <https://doi.org/10.1016/j.jcrr.2012.11.010>.
- [12] Gaudino M, Di Castelnuovo A, Zamparelli R, Andreotti F, Burzotta F, Iacoviello L, et al. Genetic control of postoperative systemic inflammatory reaction and pulmonary and renal complications after coronary artery surgery. *J Thorac Cardiovasc Surg* 2003;126(4):1107–12. Available from: <https://doi.org/10.1016/S0022>.
- [13] Grabulosa CC, Batista MC, Cendoroglo M, Quinto BM, Narciso R, Monte JC, et al. Frequency of TGF- $\beta$  and IFN- $\gamma$  genotype as risk factors for acute kidney injury and death in intensive care unit patients. *Biomed Res Int* 2014;2014:904730. Available from: <https://doi.org/10.1155/2014/904730>.
- [14] Veldman BA, Spiering W, Doevendans PA, Vervoort G, Kroon AA, de Leeuw PW, et al. The Glu298Asp polymorphism of the NOS 3 gene as a determinant of the baseline production of nitric oxide. *J Hypertens* 2002;20(10):2023–7.
- [15] Bányász I, Bokodi G, Vászárhelyi B, Treszl A, Derzbach L, Szabó A, et al. Genetic polymorphisms for vascular endothelial growth factor in perinatal complications. *Eur Cytokine Netw* 2006;17(4):266–70.
- [16] Perianayagam MC, Liangos O, Kolyada AY, Wald R, MacKinnon RW, Li L, et al. NADPH oxidase p22phox and catalase gene variants are associated with biomarkers of oxidative stress and adverse outcomes in acute renal failure. *J Am Soc Nephrol* 2007;18(1):255–63. Available from: <https://doi.org/10.1681/ASN.2006070806>.
- [17] Kornek M, Deutsch M-A, Eichhorn S, Lahm H, Wagenpfeil S, Krane M, et al. COMT-158-polymorphism is not a risk factor for acute kidney injury after cardiac surgery. *Dis Markers* 2013;35(2). Available from: <https://doi.org/10.1155/2013/279046>.

- [18] Zhao B, Lu Q, Cheng Y, Belcher JM, Siew ED, Leaf DE, et al. A genome-wide association study to identify single-nucleotide polymorphisms for acute kidney injury. *Am J Respir Crit Care Med* 2017;195(4):482–90. Available from: <https://doi.org/10.1164/rccm.201603-0518OC>.
- [19] Köttgen A. Genome-wide association studies in nephrology research. *Am J Kidney Dis* 2010;56(4):743–58. Available from: <https://doi.org/10.1053/j.ajkd.2010.05.018>.
- [20] Stafford-Smith M, Li YJ, Mathew JP, Li YW, Ji Y, Phillips-Bute BG, et al. Genome-wide association study of acute kidney injury after coronary bypass graft surgery identifies susceptibility loci. *Kidney Int* 2015;88(4):823–32. Available from: <https://doi.org/10.1038/ki.2015.161>.
- [21] Kavanagh D, Goodship TH, Richards A. Atypical hemolytic uremic syndrome. *Semin Nephrol* 2013;33(6):508–30. Available from: <https://doi.org/10.1016/j.semnephrol.2013.08.003>.
- [22] Coresh J, Selvin E, Stevens LA, Manzi J, Kusek JW, Eggers P, et al. Prevalence of chronic kidney disease in the United States. *J Am Med Assoc* 2007;298(17):2038–47. Available from: <https://doi.org/10.1001/jama.298.17.2038>.
- [23] Baum M. Overview of chronic kidney disease in children. *Curr Opin Pediatr* 2010;22(2):158–60. Available from: <https://doi.org/10.1097/MOP.0b013e32833695cb>.
- [24] Eknoyan G, Lameire N, Barsoum R, Eckardt KU, Levin A, Levin N, et al. The burden of kidney disease: improving global outcomes. *Kidney Int* 2004;66(4):1310–14. Available from: <https://doi.org/10.1111/j.1523-1755.2004.00894.x>.
- [25] Arpegård J, Viktorin A, Chang Z, de Faire U, Magnusson PK, Svensson P. Comparison of heritability of Cystatin C- and creatinine-based estimates of kidney function and their relation to heritability of cardiovascular disease. *J Am Heart Assoc* 2015;4(1):e001467. Available from: <https://doi.org/10.1161/jaha.114.001467>.
- [26] Skrunes R, Svarstad E, Reisæter AV, Vikse BE. Familial clustering of ESRD in the Norwegian population. *Clin J Am Soc Nephrol* 2014;9(10):1692–700. Available from: <https://doi.org/10.2215/CJN.01680214>.
- [27] Levey AS, Coresh J. Chronic kidney disease. *Lancet* 2012;379(9811):165–80. Available from: [https://doi.org/10.1016/S0140-6736\(11\)60178-5](https://doi.org/10.1016/S0140-6736(11)60178-5).
- [28] Smith JM, Stablein DM, Munoz R, Hebert D, McDonald RA. Contributions of the Transplant Registry: the 2006 annual report of the North American Pediatric Renal Trials and Collaborative Studies (NAPRTCS). *Pediatr Transplant* 2007;11(4):366–73. Available from: <https://doi.org/10.1111/j.1399-3046.2007.00704.x>.
- [29] Nestor JG, Groopman EE, Gharavi AG. Towards precision nephrology: the opportunities and challenges of genomic medicine. *J Nephrol* 2018;31(1):47–60. Available from: <https://doi.org/10.1007/s40620-017-0448-0>.
- [30] Prakash S, Gharavi AG. Diagnosing kidney disease in the genetic era. *Curr Opin Nephrol Hypertens* 2015;24(4):380–7. Available from: <https://doi.org/10.1097/MNH.0000000000000139>.
- [31] Savige J, Gregory M, Gross O, Kashtan C, Ding J, Flinter F. Expert guidelines for the management of Alport syndrome and thin basement membrane nephropathy. *J Am Soc Nephrol* 2013;24(3):364–75. Available from: <https://doi.org/10.1681/ASN.2012020148>.
- [32] Lovric S, Ashraf S, Tan W, Hildebrandt F. Genetic testing in steroid-resistant nephrotic syndrome: when and how? *Nephrol Dial Transplant* 2016;31(11):1802–13. Available from: <https://doi.org/10.1093/ndt/gfv355>.
- [33] Oud MM, Lamers IJ, Arts HH. Ciliopathies: genetics in pediatric medicine. *J Pediatr Genet* 2017;6(1):18–29. Available from: <https://doi.org/10.1055/s-0036-1593841>.
- [34] Hwang DY, Dworschak GC, Kohl S, Saisawat P, Vivante A, Hilger AC, et al. Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. *Kidney Int* 2014;85(6):1429–33. Available from: <https://doi.org/10.1038/ki.2013.508>.
- [35] Lemaire M, Frémeaux-Bacchi V, Schaefer F, Choi M, Tang WH, Le Quintrec M, et al. Recessive mutations in DGKE cause atypical hemolytic-uremic syndrome. *Nat Genet* 2013;45(5):531–6. Available from: <https://doi.org/10.1038/ng.2590>.
- [36] Ozaltin F, Li B, Rauhauser A, An SW, Soylemezoglu O, Gonul II, et al. DGKE variants cause a glomerular microangiopathy that mimics membranoproliferative GN. *J Am Soc Nephrol* 2013;24(3):377–84. Available from: <https://doi.org/10.1681/ASN.2012090903>.
- [37] Caruana G, Wong MN, Walker A, Heloury Y, Webb N, Johnstone L, et al. Copy-number variation associated with congenital anomalies of the kidney and urinary tract. *Pediatr Nephrol* 2015;30(3):487–95. Available from: <https://doi.org/10.1007/s00467-014-2962-9>.
- [38] Verbitsky M, Kogon AJ, Matheson M, Hooper SR, Wong CS, Warady BA, et al. Genomic disorders and neurocognitive impairment in pediatric CKD. *J Am Soc Nephrol* 2017;28(8):2303–9. Available from: <https://doi.org/10.1681/ASN.2016101108>.
- [39] Ruebner RL, Laney N, Kim JY, Hartung EA, Hooper SR, Radcliffe J, et al. Neurocognitive dysfunction in children, adolescents, and young adults with CKD. *Am J Kidney Dis* 2016;67(4):567–75. Available from: <https://doi.org/10.1053/j.ajkd.2015.08.025>.
- [40] Tzur S, Rosset S, Shemer R, Yudkovsky G, Selig S, Tarekegn A, et al. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet* 2010;128(3):345–50. Available from: <https://doi.org/10.1007/s00439-010-0861-0>.
- [41] Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 2010;329(5993):841–5. Available from: <https://doi.org/10.1126/science.1193032>.
- [42] Köttgen A, Hwang SJ, Larson MG, Van Eyk JE, Fu Q, Benjamin EJ, et al. Uromodulin levels associate with a common UMOD variant and risk for incident CKD. *J Am Soc Nephrol* 2010;21(2):337–44. Available from: <https://doi.org/10.1681/ASN.2009070725>.
- [43] Trudu M, Janas S, Lanzani C, Debaix H, Schaeffer C, Ikehata M, et al. Common noncoding UMOD gene variants induce salt-sensitive hypertension and kidney damage by increasing uromodulin expression. *Nat Med* 2013;19(12):1655–60. Available from: <https://doi.org/10.1038/nm.3384>.
- [44] Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet* 2013;14:295–300 England.
- [45] McCarthy HJ, Bierzynska A, Wherlock M, Ognjanovic M, Kerecuk L, Hegde S, et al. Simultaneous sequencing of 24 genes associated with steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol* 2013;8(4):637–48. Available from: <https://doi.org/10.2215/cjn.07200712>.



- [46] Braun DA, Lawson JA, Gee HY, Halbritter J, Shril S, Tan W, et al. Prevalence of monogenic causes in pediatric patients with nephrolithiasis or nephrocalcinosis. *Clin J Am Soc Nephrol* 2016;11(4):664–72. Available from: <https://doi.org/10.2215/cjn.07540715>.
- [47] Schueler M, Halbritter J, Phelps IG, Braun DA, Otto EA, Porath JD, et al. Large-scale targeted sequencing comparison highlights extreme genetic heterogeneity in nephronophthisis-related ciliopathies. *J Med Genet* 2016;53(3):208–14. Available from: <https://doi.org/10.1136/jmedgenet-2015-103304>.
- [48] Bullich G, Domingo-Gallego A, Vargas I, Ruiz P, Lorente-Grandoso L, Furlano M, et al. A kidney-disease gene panel allows a comprehensive genetic diagnosis of cystic and glomerular inherited kidney diseases. *Kidney Int* 2018;94(2):363–71. Available from: <https://doi.org/10.1016/j.kint.2018.02.027>.
- [49] Ashraf S, Gee HY, Woerner S, Xie LX, Vega-Warner V, Lovric S, et al. ADCK4 mutations promote steroid-resistant nephrotic syndrome through CoQ10 biosynthesis disruption. *J Clin Invest* 2013;123(12):5179–89. Available from: <https://doi.org/10.1172/JCI69000>.
- [50] Ozaltin F. Primary coenzyme Q10 (CoQ 10) deficiencies and related nephropathies. *Pediatr Nephrol* 2014;29(6):961–9. Available from: <https://doi.org/10.1007/s00467-013-2482-z>.
- [51] Montini G, Malaventura C, Salviati L. Early coenzyme Q10 supplementation in primary coenzyme Q10 deficiency. *N Engl J Med* 2008;358:2849–50.
- [52] Mesnard L, Muthukumar T, Burbach M, Li C, Shang H, Dadhania D, et al. Exome sequencing and prediction of long-term kidney allograft function. *PLoS Comput Biol* 2016;12(9). Available from: <https://doi.org/10.1371/journal.pcbi.1005088> e1005088.



# Molecular basis of chronic neurodegeneration

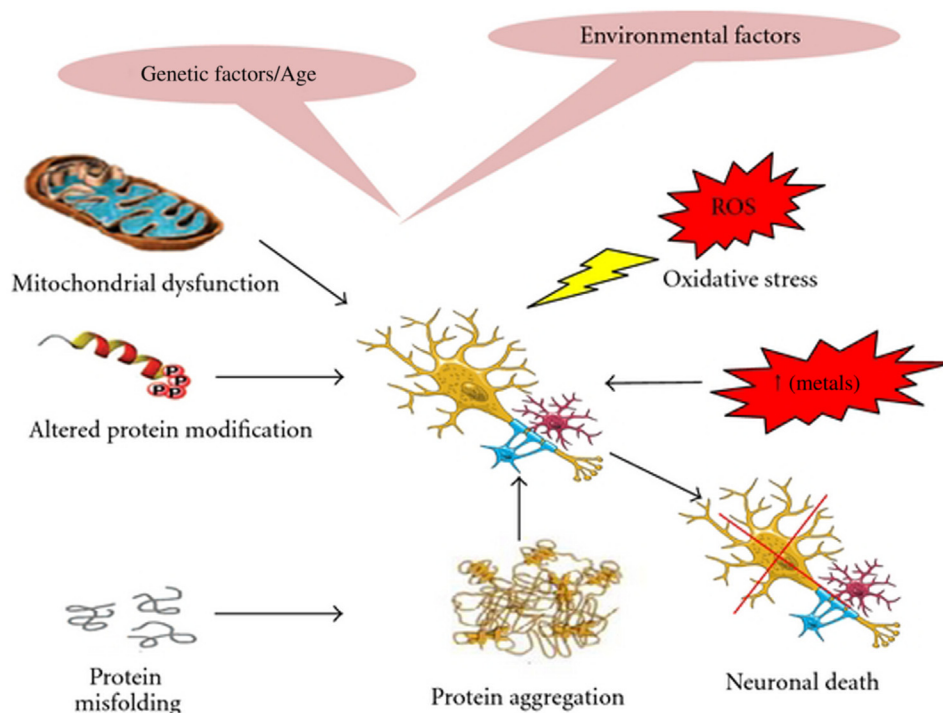
Megha Agrawal<sup>1,2</sup>

<sup>1</sup>Department of Bioengineering, University of Illinois at Chicago, Chicago, IL, United States <sup>2</sup>USA Prime Biotech LLC & Biotechnology, Kiosk Florida, United States

## 26.1 Introduction

Neurodegeneration or, more precisely, neurodegenerative disorders characterize complex and serious medical conditions, which principally affect the neurons in the human brain. Such conditions lead to the disorders of the central nervous system (CNS), which eventually results in the progressive loss of neural tissues including death of neurons [1]. Due to the inability of the neurons to self-regenerate after neurodegenerative cell death or severe damage that occurs to the neural tissue, the neurodegenerative disorders and the concomitant neurodegenerative diseases do not have natural cures on their own. Over the last decades, there is a vast array of research that has been conducted to help advance our fundamental and essential understanding of the fatal neurodegenerative diseases such as Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), and other various neurodegeneration-initiated cerebral failures. At present, roughly around 5 million Americans suffer from AD; 1 million from PD; 400,000 from multiple sclerosis; 30,000 from amyotrophic lateral sclerosis (ALS); and 30,000 from HD [1]. These alarming numbers are expected to rise significantly as the population ages due to the fact that neurodegenerative diseases strike primarily in mid-to-late life. Enormous volumes of literature suggest that these neurodegenerative disorders arise due to several factors. These include abnormal protein dynamics with defective protein degradation and aggregation; oxidative stress and free radical formation (ROS) and impaired bioenergetics and mitochondrial dysfunction and exposure to metal toxicity and pesticides. Fig. 26.1 schematically describes different factors associated with neurodegenerative disease [2].

The research conducted over the past years by neuroscientists across the globe has shed light on the limitations and boundaries to treat neurodegeneration. Particularly as the neural deterioration ensues and neurodegeneration progresses, the inability of the neurons to regenerate or regain functionality on their own poses a serious clinical challenge to effectively treat neurodegeneration. In this regard, latest research has shown that stem cell-based therapy can be potentially advantageous in neuroregeneration or even neuronal cell replacement [3,4]. One of the most important missions of effective diagnosis and prognosis of neurodegeneration lies in the ability of early detection of the onset of neurodegeneration. Timely and prompt diagnosis of the disease is critical as it offers a chance for an early treatment that may be helpful in preventing further progression of the deadly neurodegeneration and its aftermaths that takes millions of lives every year globally [5]. Among current diagnostics, neuropathologic techniques are considered one of the viable and sustainable approaches [1,6]. This conversely is usually based on an autopsy that is done after the death of a patient. Medical researchers are in the pursuit for effective noninvasive diagnostic methods that can be deployed for an early detection of neurodegeneration, when a pharmacological intervention is still possible. It has been demonstrated that molecular diagnostics can provide a powerful method to detect and diagnose various neurological diseases [2]. The confirmation of such diagnosis permits early detection and subsequent medical counseling that benefits specific patients to undergo clinically important drug trials. This offers a medical pathway to have improved insight of neurogenesis and thus leading to eventual cure of the neurodegenerative diseases.



**FIGURE 26.1** A schematic illustration of different factors associated with neurodegenerative disease. Source: Reproduced with permission from Sheikh S, et al. Neurodegenerative diseases: multifactorial conformational diseases and their therapeutic interventions. *J Neurodegener Dis* 2013;2013:563481.

## 26.2 Neurodegenerative disease clinical case studies and molecular systems underpinning the clinical scenario

Over the last decade, the field of neurobiology has progressed with reports of major significance that have provided increasing insights into the pathophysiology of neurodegenerative diseases and has thus opened new avenues to the development of molecular targeted therapies. In this context, many compounds have been tested and optimized, which have exhibited positive results in animal studies though there is almost no drug for which the efficacy has been confirmed in clinical trials [7]. The important issues related to limited knowledge of the exact mechanism of neuron loss; neurotoxic safety and safe delivery of compounds to the target is very few established animal models compatible with human pathology, and lack of better tools to diagnose presymptomatic patients make it difficult for clinical trials of drugs. Adding up to these issues, there are limited financial and patient resources. It is therefore suggested that the efficiency of both basic and clinical studies may be improved by combining integrated approaches in order to investigate the efficacy of potential disease modifying agents. It is thus an imperative part of any case study to verify the reproducibility of positive results from animal experiments that analyze the efficacy of compounds at symptomatic stages. By doing this, the viability and credibility of pre-clinical studies would improvise.

Globally, neuroresearchers are involved in developing and elaborating novel designs of phase 2 clinical trials, for example, futility study to build proof-of-concept processes. Specifically, it is believed and understood that thorough analyses of natural histories of biological and neurophysiological markers may provide critically important information for designing such preventive trials. Conceptual innovations are required in such trials to overcome limitations of conventional approaches for the development of molecular targeted therapies [7].

Furthermore, molecular systems in neurodegenerative diseases need to be studied and deliberated in order to completely understand and correlate clinical trials of neurodegenerative disorders. It is understood that most neurodegenerative disorders are associated with abnormal proteins that are accumulated in the CNS. Consequently, this accumulation of protein leads to neuronal degeneration and eventually neuronal death. Recent progressions in our understanding of molecular systems and genetics provided us both with the

possibility of identification of the abnormal proteins that are involved in many neurodegenerative disorders and also the genes that encode these proteins. These studies have directed us to a much better understanding of the underlying mechanisms of these neuro-disorders, although this understanding is still limited and inadequate due to the lack of accompanying major improvements in clinical trials and diagnostic tests or treatments for these disorders. An example to illustrate this inadequacy is the area that corresponds to the prion diseases that are rare neurodegenerative disorders and are believed to be associated with the accumulation of a misfolded host protein in the CNS. Prion diseases have been considered a paradigm for protein-misfolding diseases. However, there are significant differences between prion diseases and other neurodegenerative disorders, and it is imperative that more advances in the diagnostics and clinical studies are required to shed light on prion diseases [8].

In this section, we describe some prominent case studies related to symptoms, neuropsychological testing, and brain pathology of patients with various neurodegenerative disorders along with reflections on the molecular systems underpinning the clinical scenario.

### 26.2.1 Alzheimer's disease

AD commonly presents with impairment of memory and language functions. In a significant early case study of a patient with AD, language difficulties were noted more conspicuously than was the memory impairment. In the study involving throughout the limbic system and neocortex of the patient, large numbers of senile plaques and neurofibrillary tangles were observed that are believed to be the pathological hallmarks of AD [9].

Comprehensive neuropathologic examination was conducted that allowed a definitive diagnosis of AD. Examination of the entire brain revealed normal brain structures with certain overall shrinkage (atrophy). Fig. 26.1 shows this feature present somewhat even in an external view of the removed brain. However, atrophy was observed more prominently in a brain slice with an increase in the size of the fluid-filled cavities of the brain (the ventricles), as well as of the fissures and sulci (crevices), as can be seen in the inset of Fig. 26.2 [9].

It is worthwhile to note that in this case study with a patient with AD, stains revealed large numbers of amyloid plaques (senile plaques) and neurofibrillary tangles present throughout the limbic system and neocortex that are the pathological hallmarks of AD. Plaques are accumulations of protein outside of brain cells. In a standard histological stains test a trained pathologist can reveal plaques; however, they can be revealed more prominently with viewing a silver salt-impregnation stained section under the light microscope (Fig. 26.3) [10]. In addition to the classic and standard findings of plaques and tangles, the investigators revealed in their neuropathological examination some other abnormalities characteristic of AD pathology. These abnormalities included granulovacuolar degeneration of the pyramidal cells of the hippocampal formation (in which dying neurons contain particle-filled bubbles). Furthermore, the presence of beta-amyloid immunoreactivity in the walls of blood vessels was also observed in the patient that indicated amyloid angiopathy [9].



**FIGURE 26.2** Anatomic pathological examination of the human brain. It presents a lateral external view of the removed brain of a patient with AD. The left side shows the front of the brain (frontal lobe), while the rear (occipital lobe) side of the brain is on the right. It shows the atrophy, or loss of brain substance, which is evident in that the brain gyri are somewhat thinned, and the brain sulci and fissures are somewhat widened. The inset figure shows a coronal 1-cm section of the right hemisphere of the same brain. It shows enlargements of the fluid spaces, including the lateral ventricle (*asterisk*), Sylvian fissure (*arrow*), and sulci (*arrowheads*). AD, Alzheimer's disease. Source: *Reproduced with permission from Honig LS, Chin SS. Alzheimer's disease. Sci Aging Knowledge Environ 2001;2001(1):dn2.*



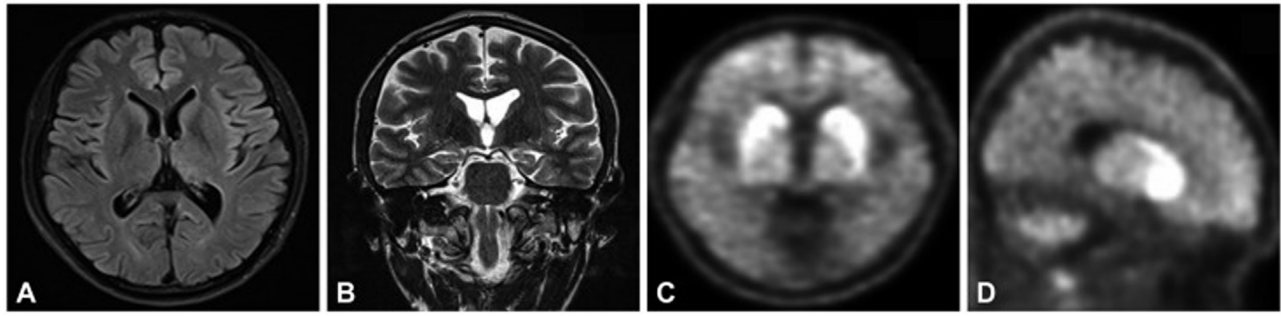


**FIGURE 26.3** A light microscope image showing amyloid plaques in a brain section viewed at low magnification. A silver-stained (modified Bielschowsky) histological section from the brain of the patient revealed many senile plaques (arrows). Source: *Reproduced with permission from Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. Lancet 2006;368(9533):387–403.*

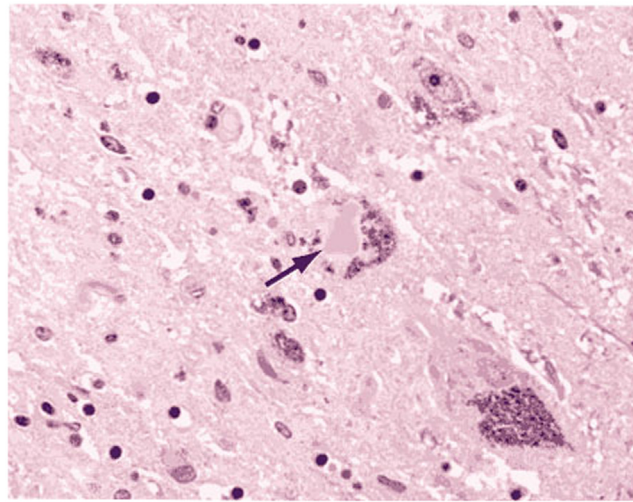
In another case study the two primary cardinal lesions—neurofibrillary tangle and the senile plaque—which are associated with AD have been corroborated [11]. The case study investigators have acknowledged that it may be very difficult to distinguish between AD, especially in its early stages and normal aging. This is especially the situation while examining the brains of the patients who died at an advanced old age. Further, the investigators also noted that in the absence of other coexistent neurodegenerative disease progressions, such as infarctions or PD-related lesions, instances of pure forms of AD are relatively uncommon. Therefore this fact must be taken into account by researchers who conduct postmortem brain tissues for research [11]. Moreover, it is widely suggested to have a correlation of AD neuropathologic changes with cognitive status. The reason is that most people in advanced old age can have significant non-AD brain lesions that may alter cognition independently of AD. It is therefore recommended that neuropathology of aging-related brain disease must take into account diverse medical, technical, biochemical, and anatomic considerations [12].

Recently, an important case study was conducted of an early-onset AD patient who received a confirmatory diagnosis of AD by beta-amyloid imaging. In this study the patient demonstrated an early onset of cognitive decline with an accelerated deterioration. There is relatively limited evidence on the clinical application of beta-amyloid imaging in early-onset AD patients, and therefore this case study is noteworthy and can contribute to our understanding of beta amyloid in patients with an early onset of AD [13].

The investigators employed brain magnetic resonance imaging (MRI) that demonstrated global cerebral atrophy of grade 1 by cortical atrophy scale and notable medial temporal lobe atrophy of grade 2 by medial temporal lobe atrophy visual rating scale, Fig. 26.4A and B. Further, an early onset of dementia symptoms shown by the patient made the patient a suitable case study for amyloid positron emission tomography (PET) imaging. 18-Florbetaben PET images revealed diffused amyloid deposition with score in brain beta-amyloid plaque load, with predominant amyloid deposition in the striatum (Fig. 26.4C and D) [13].



**FIGURE 26.4** Brain magnetic resonance imaging of an early-onset AD patient. Fluid-attenuated inversion recovery axial image (A), T2-weighted coronal image (B), and axial and sagittal images from amyloid imaging with 18F-florbetaben (C and D). AD, Alzheimer's disease. Source: Reproduced with permission from Um YH, et al. A case report of a 37-year-old Alzheimer's disease patient with prominent striatum amyloid retention. *Psychiatry Invest* 2017;14(4):521–4.



**FIGURE 26.5** A higher magnification view of the substantia nigra of the patient reveals the presence of pale, pinkish inclusion bodies in the cytoplasm of some neurons (arrow). These bodies correspond to Lewy bodies. Further, immunological staining of such sections reveals their immunoreactive nature for both ubiquitin and synuclein. Source: Reproduced with permission from Constantino AE, Honig LS. Parkinson's disease. *Sci Aging Knowledge Environ* 2001;2001(7):dn4.

### 26.2.2 Parkinson's disease

PD is commonly associated with tremor or changes in one's ability to walk or move. PD also results in other major difficulties that include rigidity of the body, slowness of movement, and postural imbalance. The degeneration of a specific population of neuronal cells in the brain stem, in a region, called the pars compacta of the substantia nigra causes these disease symptoms. In a case study of a patient with PD, neuropathology tests showed the loss of these neuronal cells and the appearance of characteristic neuronal cytoplasmic inclusions called Lewy bodies that were believed to be composed of aggregated synuclein protein [14].

An autopsy examination of the brain of the patient was performed, which provided a definitive diagnosis of PD. By slicing of the brain stem, it was evident that there was marked pallor of the ventral midbrain. Usually, the brain stem contains a blackish pigmented band, known as the substantia nigra. In this case study, microscopic examination of midbrain sections revealed marked changes in the substantia nigra. It showed significantly decreased numbers of the pigmented dopaminergic neurons consisting of melanin and located in the pars compacta of the substantia nigra. Some neurons were observed to have cytoplasmic, pale, sometimes concentric-ringed, eosinophilic inclusions termed Lewy bodies (Fig. 26.5). These inclusions are believed to consist of ubiquitinated synuclein protein present in dense aggregates. Further, from this case study, it can be concluded that

degeneration of the pigmented neurons in the substantia nigra pars compacta, together with the presence of Lewy bodies, is sufficient to diagnose pathological PD. It can further be concluded that the loss of neurons is characteristically accompanied by the presence of free melanin, which is thought to be detritus from dead cells. This is also due to the presence of gliosis, which is a reactive process including increased astrocytic cells and processes [14].

The other regions of the brain, notably the pigmented brain stem nuclei, contain Lewy bodies. They may also be found in the large cholinergic projecting neurons of the nucleus basalis of Meynert, in the substantia innominata. In various cases of PD, particularly those of long duration or in cases with dementia, Lewy bodies are also found to varying degrees in the cerebral neocortex. It can be stated that there is a pathological overlap between the clinicopathological entities of dementia with Lewy bodies and PD [14].

Clinical studies have revealed that fatigue and mental health can be crucial in patients with PD. Destruction of neurons and synapses can result in cognitive slowing and states of confusion (and all other motor and non-motor symptoms) in PD patients. Due to mental fatigue, patients with Parkinson's can feel exhausted after just a short span of time to do the amount of effort required for basic mental tasks. Further, it has been observed that a major contributing factor to the fatigue is the high comorbidity of sleep disturbances in patients with Parkinson's [15].

It is now believed that cognitive impairment is prevalent among patients with PD. It has been proven in clinical studies that patients with PD can have a variety of cognitive deficits including executive dysfunction, visuospatial impairment, and memory loss. We note here that the exact mechanism of cognitive impairment in PD is yet to be determined. It is however known that over time, there is widespread brain atrophy and brain cell death in patients with PD dementia "PDD." Researchers in their clinical trials have shown that cell death in "PDD" is largely attributed to the development of Lewy bodies. Lewy bodies are proteins that clump together for an unknown reason and then deposit in neurons causing cell damage and eventual cell death. Lewy bodies also develop in Lewy body dementia (LBD), although in LBD the cognitive changes happen first. We are yet to determine, and there is no consensus at this time as to whether PD and LBD are two distinct disease processes or variations of the same disorder [16].

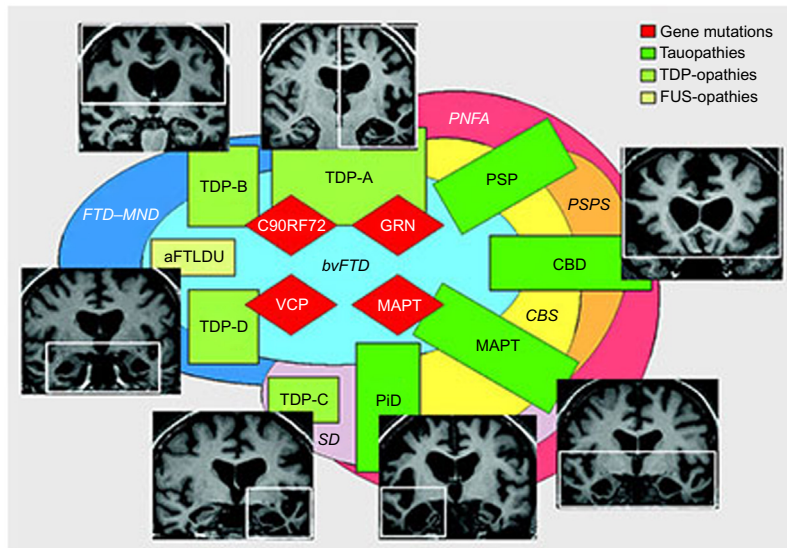
### 26.2.3 Frontotemporal dementia

Frontotemporal dementia (FTD) corresponds to either a change in personality or behavior. This includes social withdrawal, increased gregariousness, disinhibition or obsessive behaviors, or with impairment of language function. Memory losses may happen in patients with FTD. However, in the early stages of FTD, the other symptoms previously listed are more prominent than memory difficulties. Fig. 26.6 provides a schematic overview of FTD showing major clinical and neuroanatomical syndromes and diseases, and the relations between them [17].

In a particular clinical study an autopsy was ultimately performed, which revealed findings characteristic of FTD. The study discovered focal brain degeneration in the frontal and temporal regions, microscopic signs of gliosis, and cellular abnormalities of the intracellular microtubule-associated protein tau [18]. In a regionally informative pattern study, microscopic investigations of different cortical brain regions revealed specific neuropathologic findings. It unveiled the involved middle frontal, superior, middle and inferior temporal, inferior parietal, precentral, cingulate, insular, entorhinal, and transentorhinal cortices. This indicates that the anterior regions of the brain were more affected than the posterior, parietal, and occipital portions, which is a typical pattern characteristic of FTD. The research exhibited changes in neuron number and structure along with the severe loss of neurons and the presence of ballooned neuronal cells [18].

Immunostaining for a phosphorylated form of the microtubule-associated protein tau (using monoclonal antibody AT8) was also conducted. It showed diffuse abnormal intraneuronal cytoplasmic staining, sometimes referred to as "pretangles" within cortical and pyramidal neurons in the hippocampus and granular cell neurons of the dentate gyrus (gray matter under the hippocampus) [18]. It was observed that there were significant astroglial changes, including the presence of tau-positive "tuft-like" profiles in astrocytes. Further, rarefaction (loss of density) of white matter was also noted that indicated axonal loss consequent to neuronal cell loss. While some neurons were preserved, the basal ganglia and basal forebrain, gliosis (increased astrocytic cell size and number) were also found in the deep nuclei. In addition, the thalamus was found to be preserved, whereas the brain stem showed marked degeneration. There were no indications of cell inclusions such as Lewy bodies or neurofibrillary tangles; however, the substantia nigra in the midbrain exhibited prominent neuronal cell loss [18]. Moreover, there were no signs of Alzheimer's pathology in the brain, including any amyloid deposition in the form of





**FIGURE 26.6** The schematic illustration shows major genes causing frontotemporal dementia (FTD), histopathological substrates and clinical phenotypes. Neuroanatomical profiles are shown as coronal magnetic resonance imaging sections (left hemisphere displayed on the right) adjacent to the corresponding pathological substrates, with regions of predominant regional atrophy demarcated by white rectangles. Genetic bases for pathological substrates and phenotypic associations of tissue pathologies are also shown as intersecting.

#### Genes

C90RF72 = hexanucleotide repeat expansion in non-coding region of chromosome 9 open reading frame 72; GRN = progranulin; MAPT = microtubule-associated protein tau; VCP = valosin containing protein

#### Histopathologies

Diseases with cellular inclusions containing phosphorylated protein tau (tau-opathies): CBD = corticobasal degeneration; PiD = classic Pick's disease; PSP = progressive supranuclear palsy Diseases with cellular inclusions containing predominantly trans-active-response DNA binding protein 43 (TDP-opathies): histomorphological subtypes A to D (after classification scheme of Mackenzie et al<sup>2</sup>) Disease with cellular inclusions containing fused-in-sarcoma (FUS) protein (FUS-opathies): aFTL = atypical frontotemporal lobar degeneration with ubiquitinated inclusions

#### Clinical phenotypes

bvFTD = behavioral variant frontotemporal dementia; CBS = corticobasal syndrome; FTD-MND = frontotemporal dementia with motor neuron disease; PNFA = progressive non-fluent aphasia; PSPS = progressive supranuclear palsy syndrome; SD = semantic dementia. Source: Reproduced with permission from Warren JD, Rohrer JD, Rossor MN. Clinical review. Frontotemporal dementia. *Brit Med J* 2013;347:f4827.

diffuse plaques or neuritic plaques or any classical neurofibrillary tangles. The microscopic section investigations concluded regional significant and noteworthy loss of neurons and associated white matter, astroglial change, and immune histochemical evidence of tau protein pathology [18].

In another clinical case study, it was specified that FTD should be considered in middle-aged patients with insidious personality, behavioral and psychiatric changes accompanied by hyperorality and changes in diet, compulsions, stereotypical behaviors, and lack of insight. Further it was also noted that in such cases, a detailed history, physical and neurological examination, neuropsychological evaluation, and brain imaging should be performed [19].

## 26.3 Molecular pathology of neurodegenerative diseases

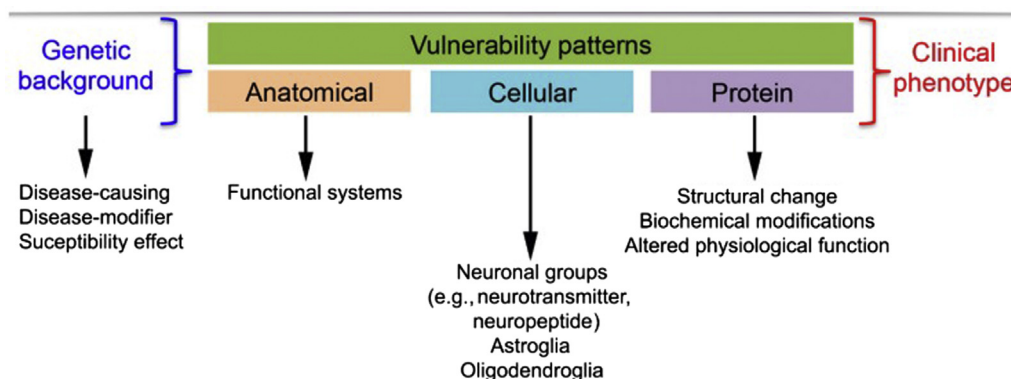
As was described in the preceding sections, neurodegenerative diseases are characterized by selective dysfunction and progressive loss of neurons. These are associated with pathologically altered proteins that not only deposit in the human brain but also in peripheral organs [20]. Despite an extensive range of studies that have been conducted to demonstrate different neurodegeneration-related proteins, such as beta amyloid, prion protein, tau,  $\alpha$ -synuclein, Transactive response (TAR) DNA-binding protein 43 (TDP-43), or fused-in-sarcoma protein, molecular classification and mechanism of neurodegenerative diseases are yet complex to comprehend. The reason for this complexity is the fact that the molecular classification relies on detailed morphological evaluation of protein deposits, their distribution in the brain, and their correlation to clinical symptoms together with specific genetic alterations. Another reason for the neuropathology-based classification is the fact that many protein deposits show a hierarchical involvement of brain regions. This has been corroborated for AD and PD and some forms of tauopathies and TDP-43 proteinopathies [20].

Classification of neurodegenerative diseases is established on the following three factors: (1) major clinical symptoms that are determined by the anatomical region showing neuronal dysfunction or loss, which may not necessarily reflect the molecular changes in the background; (2) proteins that show conformational change and biochemical modifications; and (3) cellular and subcellular pathology that corresponds to whether neurons or glial cells (either or both astro- and oligodendroglia), including which compartment of the cells, show pathological protein deposits; or whether these are found extracellularly [20]. Therefore in conjunction with a neurodegenerative syndrome, we can define anatomical, cellular and protein vulnerabilities (Fig. 26.7) [20]. We further believe that there is a possibility that genetic alterations may also lead to these alterations or influence the susceptibility to develop these neurodegenerative diseases [20].

It is important to consider anatomical involvement of neuronal loss underlying clinical symptomatology. It has been observed that in most cases, clinical symptoms do incline to overlap through the course of the disease. Therefore clinical classification can be helpful in focusing on the early symptoms. For example, cognitive decline, dementia, and alterations in high-order brain functions can be associated with involvement of the entorhinal cortex, hippocampus, limbic system, and neocortical areas. In movement disorders the basal ganglia, thalamus, brain stem nuclei, cerebellar cortex and nuclei, motor cortical areas, and lower motor neurons of the spinal cord are involved. Hence, combinations of these symptoms are observed in some diseases (i.e., prion diseases) early during the clinical course and in many neurodegenerative disorders during the progression [20].

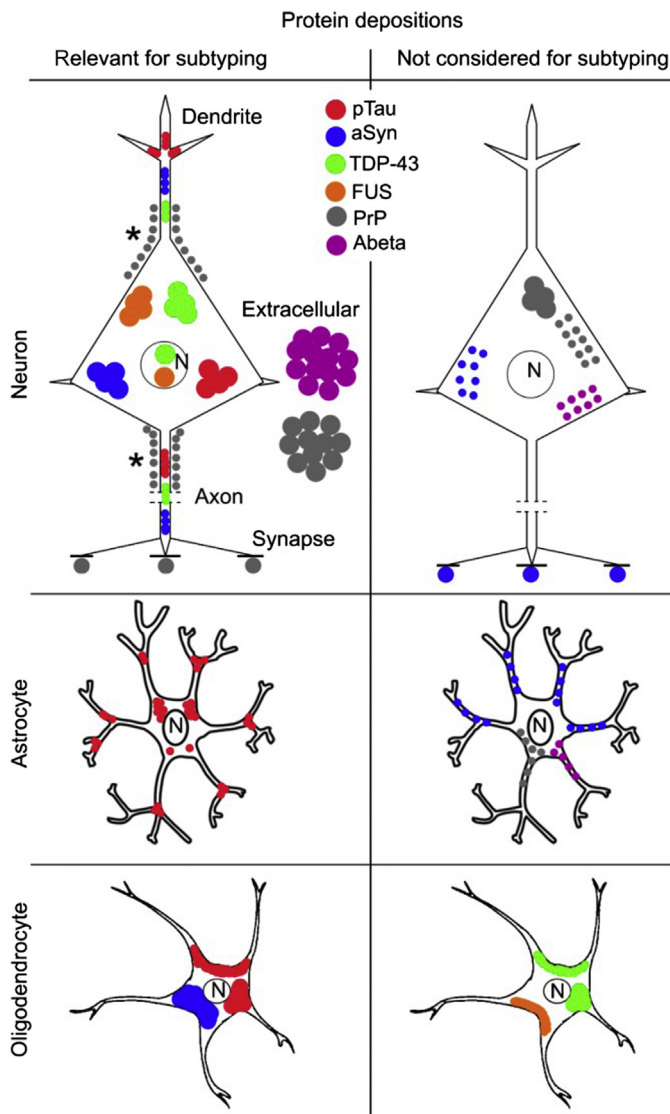
Figs. 26.8 and 26.9 present overviews of how different proteins distribute in cells and subcellular structures. This helps us understand the overlaps in the pathogenesis of different neurodegenerative disorders. These aspects associate with different biochemical modifications of a specific protein under recognized changes of the cells (i.e., lipid and metabolomics), which could have biomedical implications defining early (presymptomatic) and late (symptomatic for subtyping for prognostics) biomarkers of neurodegenerative disease [20].

In addition to the consequences of protein folding and misfolding on the neurodegeneration, mitochondrial dysfunction and reactive oxygen species, "ROS," are considered contributing factors in the progression of neurodegenerative diseases. Small oligomeric aggregates and amyloid oligomers have been extensively reported to permeabilize both cell and mitochondrial membranes. It is believed that they are responsible for calcium dysregulation, membrane depolarization, and impairment of mitochondrial functions, which have been identified as a further common feature of most neurodegenerative disorders. It is now well known that neurodegenerative diseases are a consequence of genetic mutations and/or environmental factors, which are strongly related with age [21,22]. Mitochondrial dysfunction and oxidative stress also play a pivotal role for the development of more common neurodegenerative disorders. Loss of mitochondrial function is associated with an increase in the generation of reactive oxygen intermediates [23]. It has been observed that mitochondria under normal physiological cellular conditions are intimately involved in the production of ROS through one-electron carriers in the respiratory chain, and it probably produces one to two orders lower amounts of ROS. Neuronal tissue is particularly sensitive to oxidative stress and imbalance in prooxidant versus antioxidant homeostasis in CNS results in the production of several potentially toxic ROS. These include both the radical and nonradical species that participate in the initiation and/or propagation of radical chain reactions. In AD, PD, HD, and ALS, oxidative damage is found in every class of biological molecules within neurons, spanning from lipids to DNA and proteins [2].

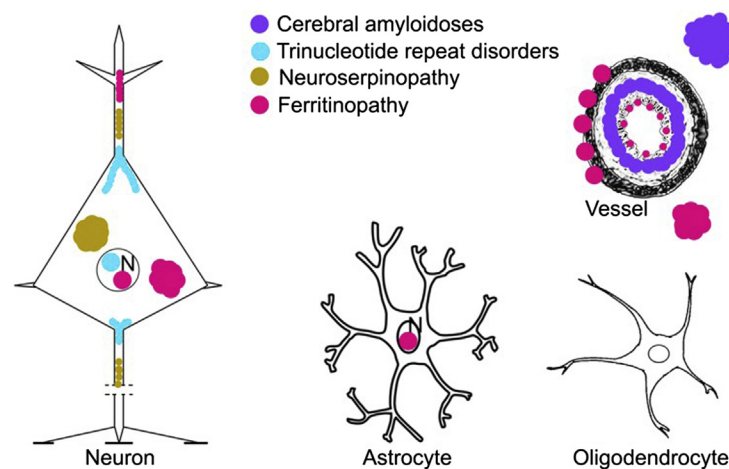


**FIGURE 26.7** Summary of the concept of vulnerability patterns in neurodegenerative diseases. Colored boxes represent different vulnerability patterns. Source: Reproduced with permission from Kovacs GG. Molecular pathological classification of neurodegenerative diseases: turning towards precision medicine. *Int J Mol Sci E* 189, 2016;17(2):1–33.





**FIGURE 26.8** Overview of cellular vulnerability in the most frequent neurodegenerative proteinopathies. Asterisk (\*) for PrP indicates periaxonal or perineuronal. Synapses are indicated by short black and bold lines (-). N, Nucleus. Source: Reproduced with permission from Kovacs GG. *Molecular pathological classification of neurodegenerative diseases: turning towards precision medicine*. *Int J Mol Sci E* 189, 2016;17(2):1–33.



**FIGURE 26.9** Cellular vulnerability patterns in rare hereditary forms of neurodegenerative diseases. Synapses are indicated by short black and bold lines (-). N, Nucleus. Source: Reproduced with permission from Kovacs GG. *Molecular pathological classification of neurodegenerative diseases: turning towards precision medicine*. *Int J Mol Sci E* 189, 2016;17(2):1–33.

## 26.4 Application of molecular diagnostics in neurodegeneration

Molecular diagnosis is a powerful tool that has emerged to be helpful in an early detection of numerous neurodegenerative disorders. One of the powerful molecular diagnostics is the application of biomarkers. Biomarkers are fundamentally biological molecular substances that are used to indicate the existence or onset of a certain disorder. Biomarkers can be used to detect both normal and abnormal biological processes. The principal requirement for a good biomarker lies in its preciseness and reliability. It should also be able to distinguish between the healthy and the diseased tissues and should at the same time differentiate among various diseases. Biomarkers are considered promising in aiding early diagnosis, and this is already setting standards for the development and improvement of new remedies to treat neuronal disorders [3,24,25]. Biomarkers may be measured using imaging techniques such as PET, MRI, and nuclear magnetic resonance spectroscopy (MRS). The research field of molecular diagnostics in neurodegenerative disorder is still a nascent area of research and development. It is anticipated that further advancements in various molecular diagnostics would pave the way for the early detection and effective treatment of neurodegeneration [1].

Table 26.1 summarizes various molecular diagnostic markers for neurodegenerative diseases [1].

**TABLE 26.1** A summary giving genetic and biochemical diagnostic markers for Alzheimer's, Parkinson's, amyotrophic lateral sclerosis, and Huntington's diseases.

Diseases	Genetic diagnostic markers	Biochemical diagnostic markers
AD	Amyloid precursor protein mutations	Plasma/CSF A $\beta_{1-42}$ peptide
	Presenilin-1 gene mutations	CSF-tau protein
	Presenilin-2 gene mutations	Phospho-tau
	<i>ApoE</i> isoforms	
	<i>ApoE</i> polymorphisms	
PD	$\alpha$ -Synuclein gene mutations	Loss of DAT
	Parkin gene mutations	Lewy bodies
	<i>UCH-L1</i> gene mutations	
	<i>PINK1</i> gene mutations	
	<i>DJ-1</i> gene mutations	
	<i>NR4A2</i> gene mutations	
ALS	<i>ALS2</i> gene mutations	mGLUR2
	<i>NEFH</i> gene mutations	SOD1
	<i>SOD1</i> gene mutation	Glutathione
	<i>C9orf72</i> gene mutation	8OH2'dG
	<i>FUS</i> gene mutation	Cytokines
	<i>TARDBP</i> gene mutations	
HD	<i>HTT</i> gene mutations	Growth hormones
		Cytokines
		mGLUR2
		SOD1
		Glutathione
		8OH2'dG

AD, Alzheimer's disease; ALS, amyotrophic lateral sclerosis; CSF, cerebrospinal fluid; DAT, dopamine transporter; HD, Huntington's disease; PD, Parkinson's disease.

Reproduced with permission from Agrawal M, Biswas A. Molecular diagnostics of neurodegenerative disorders. *Front Mol Biosci* 2015;2:54.

Biomarkers combined with MRS have shown promising results among other recently developed molecular diagnostics. It is a quantitative imaging technique that allows in vivo measurement of certain neuronal metabolites as biomarkers that can be used to study metabolic dysfunctions and irreversible neuronal damage [26]. The potential role of MRS as an in vivo molecular imaging biomarker was investigated for early diagnosis of PD and for monitoring the efficacy of therapeutic interventions [26].

Altered expression of microRNAs (miRNAs) in many disease states, including neurodegeneration along with applications of miRNAs in biological fluids in different pathologies, has made them promising candidates as neurodegenerative disease biomarkers that may eventually lead to identification of new therapeutic targets [27]. The role of miRNAs in the pathogenesis of neurodegeneration was scrutinized to advance insights into the possibilities and challenges of using these small RNA molecules as a signature for neurodegenerative conditions [27]. It is known that miRNAs can be transported by exosomes that are small membrane-derived vesicles secreted by many cell types, including neurons, astrocytes, oligodendrocytes, and microglia [28,29]. Exosomes shed from stimulated blood cells and the vascular endothelium are involved in neurological disorders [30]. Their molecular composition reflects the physiological or pathophysiological changes in their cell of origin, which gives them a significant potential as biomarkers for disease diagnosis [31]. Furthermore, they can be isolated from biofluids such as blood and urine, making them very attractive targets for diagnostic application. It has been reported that exosomal amyloid peptides accumulate in the brain plaques of AD patients [32] and tau phosphorylated at Thr-181, an established biomarker for AD, is present at elevated levels in exosomes isolated from cerebrospinal fluid (CSF) specimens of AD patients with mild symptoms [33]. Studies have also revealed release of  $\alpha$ -synuclein in exosomes in an in vitro model system of PD [34]. These exosomal proteins may have great potential in clinical diagnostics and should be further explored, as the concept is still new in the biomarker discovery arena [5].

Behavioral symptoms may also be utilized for the premortem diagnosis of neurodegenerative disorders. On the other hand, the major drawback of behavioral symptoms-based diagnosis is its limitations to identify patients early in the course of their disease, when the pharmacological intervention can significantly prevent further progression of the disease, if detected early. For example, well-established behavioral tests, such as the ADAS-Cog, which are regarded as the “gold standard” for AD diagnosis, may give false-negative results for patients with mild symptoms [35].

To overcome these diagnostics challenges, current neuropathologic methods are being combined with molecular biology techniques. This has led to increased understanding of neurodegenerative disorders along with biologically based classifications of these disorders. Molecular diagnostics provides a powerful tool in the diagnosis of many neurological diseases. For example, genetic testing of mutations in disease-causing genes has been leveraged to define and classify many of the heterogeneous inherited neurodegenerative syndromes [36,37]. Changes in pathologies, biochemistries, and genetics of patients can give us comprehensive information regarding the nature of a particular disease. However, molecular testing may only be performed after cautious consideration and genetic counseling [1].

Individuals with HD overexpressed the gene, H2A histone family, and member Y (*H2AFY*) in their blood [38]. The overexpression of this gene in both the blood and the brain was validated in samples from clinical studies. Specifically, the research demonstrates a 1.6-fold overexpression of *H2AFY* in patients with HD. Recently a study showed gene expression profiling with the help of next-generation sequencing and Fluidigm technologies. This study yielded a set of five genes as potential HD biomarkers that are highly expressed in HD blood [39]. Prokineticin 2 (*PROK2*) has been suggested to have a role in the circadian rhythms alterations that have been shown to correlate with cognitive impairment in HD [40]. Pharmacological imposition of sleep slows down cognitive decline and reverses deregulation of *PROK2* in HD models. *PROK2* is proving to be a very promising biomarker of HD progression. Evidences suggest that gene repression mechanisms are also associated with HD, and zinc finger protein 238 (*ZNF238*) is a transcriptional repressor, which is involved in brain development and myogenesis [41]. A recent gene expression study showed that the increase in mRNA levels of Aquaporin 9 (*AQP9*) and presence of *AQP9* in blood could represent peripheral or central inflammatory events when accompanied with increase in levels of four other genes [42]. Annexin A3 (*ANXA3*) and cysteine-rich, transmembrane module (*CYSTM*) are two other potential biomarkers for HD [43,44]. *ANXA3* is found to be upregulated in neuronal injury models, and *CYSTM* is involved in stress response specifically heavy metal tolerance [1].

Variability in clinical phenotype of HD and potential confounds of environmental and pharmacological factors results in the use of combination of different biomarkers that might be efficient in tracking the progression of HD. Discovery of disrupted homeostasis in HD has led to the identification of various potential biomarkers. In a recent study, researchers, with the help of cross-sectional MRS, have distinguished putaminal metabolites in premanifest and early HD individuals from controls [45]. It was found that the total *N*-acetyl aspartate

(tNAA) is lower in early HD and premanifest HD than in controls, whereas the gliosis marker myoinositol was robustly elevated in early HD. In another study, it has been demonstrated that metabolite changes in the caudate nucleus and putamen of HD gene carriers around disease onset [46]. These correlations of tNAA with disease burden score suggest that this metabolite may be useful in identifying neurochemical responses to therapeutic agents.

Vasopressin has a role in fluid balance homeostasis, increase in serum concentrations of vasopressin have been reported in HD [47]. Increased concentrations of 8-hydroxy-2-deoxyguanosine, an indicator of oxidative DNA injury and increased concentrations of plasma lipid peroxide, lactic acid, 4-hydroxynoneal, and malondialdehyde in patients with HD make them a potential biomarker [48]. Decrease in glutathione peroxidase and copper–zinc superoxide dismutase was observed in erythrocytes from HD patients compared with controls. Postmortem brain and plasma samples of patients with HD have shown elevated cytokines levels including interleukins 4, 6, 8, 10, and 23; TNF- $\alpha$ ; and clusterin. The inflammatory profile differences between control and gene carriers serve as potential biochemical marker for HD including rest of the abovementioned biomarkers. All these biomarkers would facilitate accurate evaluation and assessment of the effectiveness of new therapies and therefore improvise the safety and effectiveness of clinical trials [1].

The levels of tau protein and A $\beta$  in CSF are the two most promising biochemical markers of AD. A $\beta$  is secreted into the extracellular space and biological fluids, including CSF making A $\beta$ 42 a considerable indicator of AD [49]. A decrease in levels of A $\beta$  in CSF reflects AD, and its sensitivity is around 80%–90%. As AD progresses A $\beta$  peptide from CSF aggregates to form plaques in the brain, thereby, lowering its concentration in the CSF. CSF-A $\beta$ 42 appears to be a remarkable biomarker for diagnosis of AD when used in combination with other AD biomarkers. CSF-tau also provides a very high sensitivity for AD; however, the reason for its abnormal increase in AD patients is yet not clear. The combination of both CSF-A $\beta$ 42 and CSF-tau may improve their specificity and sensitivity and thus can be an ideal biochemical marker set for AD [50].

Genes and gene products have been identified by characterizing the monogenetic autosomal-dominant forms of PD. Several gene products of the mutated genes in the autosomal-dominant forms have been linked to mitochondrial dysfunction, oxidative stress, and mishandling of impaired or aberrant forms of the gene products (e.g., oligomeric  $\alpha$ -synuclein) [5]. More than 70 mutations on the large parkin gene have been associated with the early-onset form of parkinsonism. Mutations in the parkin gene may account for PD in as many as 50% of familial cases of autosomal-recessive juvenile parkinsonism [51].

## 26.5 Summary

Neurodegenerative disorders, which are both chronic and progressive, are characterized by selective and symmetric loss of neurons in motor, sensory, or cognitive systems. Cellular and molecular mechanisms in the pathogenesis of different neurodegenerative disorders represent a multidisciplinary research area where a robust collaboration between neurologists, psychologists, biologist and biomaterials scientists, and other trained personnel with the necessary experience in managing the diseases is required. Currently, there is no cure for these disorders, and it is therefore of paramount importance that research efforts are continued to focus on understanding the molecular underpinnings behind these disorders and subsequent various neurodegenerative diseases. This will enable development of better symptom-directed therapeutics and perhaps even curative treatments. It is commonly accepted that AD is a synaptopathy meaning that there is a loss or damage of synapses. This damage to synapses leads to altered neuronal circuitry. Further, ongoing research on the molecular events underlying PD has contributed and advanced our understanding on how the normal function of specific protein associated with PD can help shed light on the causes of familial and sporadic PD.

It is believed that neurodegenerative diseases share a common pathogenetic mechanism involving aggregation and deposition of misfolded proteins, leading to inevitable progressive deterioration of CNS. It is thought-provoking to note that while the type of aggregated protein and the regional and cellular distribution of deposition may vary from one neurodegenerative disease to another, these disorders may all be linked by similar pathways of protein aggregation with fibril formation and amyloid deposition.

On the molecular diagnostics front, extensive researches have demonstrated that the application of innovative extensive approaches in the molecular analysis such as massive high-throughput sequencing (DNA and RNA) and proteomics is promising, which paves the path for the involvement of multiple cellular pathways in a given

pathogenic process. Studies have revealed that misregulated miRNAs (miRNAs: endogenous small RNAs) binding the target sites of protein-coding genes lead to their degradation or the repression of translation. This implies to suggest that alterations in miRNA regulatory pathways may contribute significantly to neurodegenerative disorders pathogenesis.

Future research directions on molecular diagnostics may include designing and developing a combination of several biomolecular diagnostic markers for multifunctionalities. Such a multifunctional molecular diagnostic technology platform would significantly augment the accuracy, specificity, and sensitivity. Developing molecular diagnostics based on circulating miRNAs could also be a highly promising approach for developing minimally invasive screening tests for neurodegenerative disorders.

Future studies may also include developing a multicenter and prospective design of molecular diagnostics tools, measurement of multiple potential biomarkers, and a prolonged clinical follow-up period (till death as end point). This would provide assessment of both clinical features and determinations of the biological diagnostics and eventually neuropathological confirmation by examining the brains of patients at death.

## References

- [1] Agrawal M, Biswas A. Molecular diagnostics of neurodegenerative disorders. *Front Mol Biosci* 2015;2:54.
- [2] Sheikh S, et al. Neurodegenerative diseases: multifactorial conformational diseases and their therapeutic interventions. *J Neurodegener Dis* 2013;2013:563481.
- [3] Chung S, et al. Genetic engineering of mouse embryonic stem cells by Nurr1 enhances differentiation and maturation into dopaminergic neurons. *Eur J Neurosci* 2002;16(10):1829–38.
- [4] Rachakonda V, Pan TH, Le WD. Biomarkers of neurodegenerative disorders: how good are they? *Cell Res* 2004;14(5):347–58.
- [5] Miller DB, O'Callaghan JP. Biomarkers of Parkinson's disease: present and future. *Metabolism* 2015;64(3 Suppl 1):S40–6.
- [6] Agrawal M, Cho WC. Molecular diagnostics in the detection of neurodegenerative disorders. Lausanne: Frontiers Media; 2017. Available from: <https://doi.org/10.3389/978-2-88945-194-4>.
- [7] Katsuno M, Tanaka F, Sobue G. Perspectives on molecular targeted therapies and clinical trials for neurodegenerative diseases. *J Neurol Neurosurg Psychiatry* 2012;83(3):329–35.
- [8] Peden AH, Ironside JW. Molecular pathology in neurodegenerative diseases. *Curr Drug Targets* 2012;13(12):1548–59.
- [9] Honig LS, Chin SS. Alzheimer's disease. *Sci Aging Knowledge Environ* 2001;2001(1):dn2.
- [10] Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. *Lancet* 2006;368(9533):387–403.
- [11] Perl DP. Neuropathology of Alzheimer's disease. *Mt Sinai J Med* 2010;77(1):32–42.
- [12] Nelson PT, et al. Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. *J Neuropathol Exp Neurol* 2012;71(5):362–81.
- [13] Um YH, et al. A case report of a 37-year-old Alzheimer's disease patient with prominent striatum amyloid retention. *Psychiatry Invest* 2017;14(4):521–4.
- [14] Constantino AE, Honig LS. Parkinson's disease. *Sci Aging Knowledge Environ* 2001;2001(7):dn4.
- [15] Thurman SA. Case study: the Parkinson's experience. PURE Insights. Western Oregon University; 2013.
- [16] Dancis A, Cotter VT. Diagnosis and management of cognitive impairment in Parkinson's disease. *Jnp J Nurse Pract* 2015;11(3):307–13.
- [17] Warren JD, Rohrer JD, Rossor MN. Clinical review. Frontotemporal dementia. *Brit Med J* 2013;347:f4827.
- [18] Honig LS, Bell K, Chin SS. Frontotemporal dementia. *Sci Aging Knowledge Environ* 2003;2003(13):DN1.
- [19] Kuğu N, Doğan O, Kavakcı O, Terlemez I. Frontotemporal dementia: a case presentation. *J Psychiatry Neurol Sci* 2010;23:293–9.
- [20] Kovacs GG. Molecular pathological classification of neurodegenerative diseases: turning towards precision medicine. *Int J Mol Sci E* 189 2011;17(2):1–33.
- [21] Mandemakers W, Morais VA, De Strooper B. A cell biological perspective on mitochondrial dysfunction in Parkinson disease and other neurodegenerative diseases. *J Cell Sci* 2007;120(Pt 10):1707–16.
- [22] Jellinger KA. Recent advances in our understanding of neurodegeneration. *J Neural Transm (Vienna)* 2009;116(9):1111–62.
- [23] Van Houten B, Woshner V, Santos JH. Role of mitochondrial DNA in toxic responses to oxidative stress. *DNA Repair (Amst)* 2006;5(2):145–52.
- [24] Peterson DA. Stem cells in brain plasticity and repair. *Curr Opin Pharmacol* 2002;2(1):34–42.
- [25] Agrawal M, Cho WC. Editorial: molecular diagnostics in the detection of neurodegenerative disorders. *Front Mol Biosci* 2017;4:10.
- [26] Ciurleo R, et al. Magnetic resonance spectroscopy: an in vivo molecular imaging biomarker for Parkinson's disease? *Biomed Res Int* 2014;2014:519816.
- [27] Grasso M, et al. Circulating miRNAs as biomarkers for neurodegenerative disorders. *Molecules* 2014;19(5):6891–910.
- [28] Lachenal G, et al. Release of exosomes from differentiated neurons and its regulation by synaptic glutamatergic activity. *Mol Cell Neurosci* 2011;46(2):409–18.
- [29] Russo I, Bubacco L, Greggio E. Exosomes-associated neurodegeneration and progression of Parkinson's disease. *Am J Neurodegener Dis* 2012;1(3):217–25.
- [30] Soreq L, et al. Small RNA sequencing-microarray analyses in Parkinson leukocytes reveal deep brain stimulation-induced splicing changes that classify brain region transcriptomes. *Front Mol Neurosci* 2013;6:10.
- [31] Lin J, et al. Exosomes: novel biomarkers for clinical diagnosis. *ScientificWorldJournal* 2015;2015:657086.



- [32] Rajendran L, et al. Alzheimer's disease beta-amyloid peptides are released in association with exosomes. *Proc Natl Acad Sci USA* 2006;103(30):11172–7.
- [33] Saman S, et al. Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease. *J Biol Chem* 2012;287(6):3842–9.
- [34] Alvarez-Erviti L, et al. Lysosomal dysfunction increases exosome-mediated alpha-synuclein release and transmission. *Neurobiol Dis* 2011;42(3):360–7.
- [35] Posner HB, et al. Establishing the psychometric underpinning of cognition measures for clinical trials of Alzheimer's disease and its precursors: a new approach. *Alzheimers Dement* 2013;9(1 Suppl):S56–60.
- [36] Gasser T, et al. EFNS task force on molecular diagnosis of neurologic disorders: guidelines for the molecular diagnosis of inherited neurologic diseases. Second of two parts. *Eur J Neurol* 2001;8(5):407–24.
- [37] Gasser T, et al. EFNS task force on molecular diagnosis of neurologic disorders: guidelines for the molecular diagnosis of inherited neurologic diseases. First of two parts. *Eur J Neurol* 2001;8(4):299–314.
- [38] Hu Y, et al. Transcriptional modulator H2A histone family, member Y (H2AFY) marks Huntington disease activity in man and mouse. *Proc Natl Acad Sci USA* 2011;108(41):17141–6.
- [39] Mastrokolias A, et al. Huntington's disease biomarker progression profile identified by transcriptome sequencing in peripheral blood. *Eur J Hum Genet* 2015;23(10):1349–56.
- [40] Aziz NA, et al. Sleep and circadian rhythm alterations correlate with depression and cognitive impairment in Huntington's disease. *Parkinsonism Relat Disord* 2010;16(5):345–50.
- [41] Zhai W, et al. In vitro analysis of huntingtin-mediated transcriptional repression reveals multiple transcription factor targets. *Cell* 2005;123(7):1241–53.
- [42] Mesko B, et al. Peripheral blood gene expression patterns discriminate among chronic inflammatory diseases and healthy controls and identify novel targets. *BMC Med Genomics* 2010;3:15.
- [43] Borovecki F, et al. Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. *Proc Natl Acad Sci USA* 2005;102(31):11023–8.
- [44] Venancio TM, Aravind L. CYSTM, a novel cysteine-rich transmembrane module with a role in stress tolerance across eukaryotes. *Bioinformatics* 2010;26(2):149–52.
- [45] Sturrock A, et al. A longitudinal study of magnetic resonance spectroscopy Huntington's disease biomarkers. *Mov Disord* 2015;30(3):393–401.
- [46] van den Bogaard SJ, et al. Longitudinal metabolite changes in Huntington's disease during disease onset. *J Huntingtons Dis* 2014;3(4):377–86.
- [47] Wood NI, et al. Increased thirst and drinking in Huntington's disease and the R6/2 mouse. *Brain Res Bull* 2008;76(1–2):70–9.
- [48] Weir DW, Sturrock A, Leavitt BR. Development of biomarkers for Huntington's disease. *Lancet Neurol* 2011;10(6):573–90.
- [49] Sunderland T, et al. Decreased beta-amyloid1-42 and increased tau levels in cerebrospinal fluid of patients with Alzheimer disease. *J Am Med Assoc* 2003;289(16):2094–103.
- [50] Sjogren M, Andreasen N, Blennow K. Advances in the detection of Alzheimer's disease-use of cerebrospinal fluid biomarkers. *Clin Chim Acta* 2003;332(1–2):1–10.
- [51] Pankratz N, et al. Significant linkage of Parkinson disease to chromosome 2q36-37. *Am J Hum Genet* 2003;72(4):1053–7.

# Molecular basis of movement disorders

*Claire MacIver, Dafydd Morgan and Kathryn J. Peall*

MRC Neuropsychiatric Genomics Institute, Cardiff University, Cardiff, United Kingdom

## 27.1 Introduction

Movement disorders form a broad phenomenological spectrum of conditions including tremor, myoclonus, dystonia, chorea, parkinsonism, and ataxia, as well as spanning both neurodevelopmental and neurodegenerative fields. Huntington's disease (HD), a neurodegenerative choreiform disorder, was one of the first movement disorders for which an underlying genetic cause was identified. Following this, disease-causing gene discovery predominantly focused on linkage analysis in large, multigenerational families where multiple family members were affected with the same disorder. However, it is the advent of next-generation sequencing that has led to the rise in disease-causing genes being identified for movement disorders, which in some cases has facilitated further understanding of disease mechanisms and provided insights for future therapeutic development. In this chapter, we will discuss the genetically determined forms of parkinsonism, dystonia, ataxia, and other movement disorder types identified to date as well as the likely mechanisms by which mutations in these genes lead to the disorder phenotype.

## 27.2 Parkinson's disease

A summary of the disease-causing genes in Parkinson's and their associated clinical phenotypes can be summarized in [Table 27.1](#).

### 27.2.1 Epidemiology and pathophysiology

Parkinson's disease is characterized by alpha-synuclein deposition and dopaminergic neuronal degeneration in the substantia nigra. The classical clinical signs of parkinsonism include bradykinesia, a resting tremor, and rigidity. Parkinson's is age-related with 1% of the population affected at 65 years of age, increasing to 4%–5% by the age of 85 [1]. Although the cause of Parkinson's disease is unknown, there seems to be an interplay between genetic and environmental factors. Environmental factors associated with an increased risk of Parkinson's include pesticide use, rural living, well-water consumption, and occupational exposure, for example, mining. Protective agents against the development of Parkinson's disease include smoking, alcohol consumption, and caffeine intake [1]. Six genes, namely, *SNCA*, *LRRK2*, *Parkin*, *PINK1*, *DJ1*, and *VPS35*, with Mendelian patterns of inheritance have been identified as causing Parkinson's disease when carrying pathogenic mutations.

### 27.2.2 Genetics

#### 27.2.2.1 Autosomal dominant inheritance

##### 27.2.2.1.1 SNCA/PARK1: alpha-synuclein gene

Located on chromosome 4q21-22, the *SNCA* gene gives rise to the alpha-synuclein protein. Little is known about the normal function of *SNCA*; however, roles in regulating the release of neurotransmitters, synaptic

**TABLE 27.1** Genetic and clinical characteristics of Parkinson's disease.

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype	
						Motor	Nonmotor
168601/605543	2.5% of unrelated affected carriers	AD	SNCA/ PARK1/4 (4q22.1)	Alpha-synuclein	Toxicity caused by aggregates of mutated protein	Parkinsonism Dysarthria	Dementia
607060	0.5%–2% sporadic cases and 5% familial cases in Caucasians. 18%–30% in Ashkenazi Jews/North African Barbar	AD	LRRK2/ PARK8 (12q12)	Leucine-rich repeat serine/threonine-protein kinase 2	Impaired autophagy	Tremor-dominant parkinsonism, dystonia	Depression
614203	Rare	AD	VPS35/ PARK17 (16q11.2)	Vacuolar protein sorting-associated protein 35	Protein aggregation, mitochondrial dysfunction, increased ROS production	Parkinsonism, dystonia, dyskinesia	Memory impairment
600116	2.5%–8.2% <45 years	AR	Parkin/ PARK2 (6q26)	E3 ubiquitin-protein ligase parkin	Impaired mitophagy	Parkinsonism, early-onset dystonia, hyperreflexia, slower progression of disease	Depression, constipation
605909	1%–2% of early-onset	AR	PINK1/ PARK6 (1p36.12)	Serine/threonine-protein kinase PINK1	Impaired mitophagy	Early-onset parkinsonism	Depression
606324	<1% of early-onset parkinsonism	AR	DJ1/ PARK7 (1p36.23)	Protein/nucleic acid deglycase DJ-1	Exposure to oxidative stress causing neurotoxicity	Early-onset parkinsonism, slower progression	Psychosis

AD, Autosomal dominant; AR, autosomal recessive; ROS, reactive oxygen species.

function, and neural plasticity have been suggested. The majority of mutations are missense mutations, with less frequent forms including gene multiplication (duplication or triplication). *SNCA* mutations result in increased expression and/or aggregation of alpha-synuclein with formation of oligomers, fibrils, and protofibrils thought to be toxic to glia, implicated in mitochondrial damage and affecting membrane permeability [2].

#### 27.2.2.1.2 LRRK2/PARK8: leucine-rich repeat kinase 2

##### Case report

A 70-year old gentleman of North African descent presented with a 5-year history of disturbed sleep involving acting out of vivid dreams, vocalizations and involuntary limb movements. Over the same time period, he also noticed a deterioration in his sense of smell. Three years later, he became aware of an asymmetric upper limb tremor, predominantly involving his right hand, most evident when sitting quietly. He had also noticed increasing difficulty with fine motor tasks such as doing buttons, zips, and cleaning his teeth. Over the same time period his family described his mobility as having generally slowed, with a tendency to fall backward. They also described him as being lower in mood than would normally be typical for him, and some short-term memory difficulties were becoming increasingly evident in the family home. Following initial review with a neurologist he was started on levodopa (62.5 mg tds), noting an improvement to his dexterity and mobility, which further increased with higher doses of the medication.

*LRRK2* mutations are the most common type of disease-causing parkinsonian mutations in Caucasian populations, accounting for 0.5%–2% of sporadic cases and 5% of familial cases, rising to 18%–30% of cases in Ashkenazi Jewish and North African Barbar populations [1,3]. *LRRK2* encodes a 2527 amino acid protein, with the most common pathogenic mutation being the missense G2019S (p. Gly2019Ser) mutation. The protein is expressed in neurons, astrocytes, and microglia, but expression is low in the dopaminergic neurons of the substantia nigra. Although a multidomain protein, mutations generally tend to be found in the central region of the

protein, affecting the Ras of complex (Roc) GTPase protein domain and kinase domain, as well as the carboxy-terminal of Roc sequence. However, the mechanisms by which these changes give rise to pathogenicity remains poorly understood.

#### 27.2.2.1.3 VPS35/PARK17

A rare, highly penetrant, cause of autosomal-dominant Parkinson's disease is the VPS35 D620N mutation. A component of a retromer complex (membrane-associated protein), VPS35 mediates retrieval of membrane proteins via endosomes and the Golgi apparatus. Mutations in this gene impair autophagy, which may manifest as protein aggregation, mitochondrial abnormality formation, increased production of reactive oxygen species, and enhanced susceptibility to cell death [4].

### 27.2.3 Autosomal recessive inheritance

#### 27.2.3.1 *Parkin/PARK2*

##### *Case report*

*A 50-year old woman was reviewed in the neurology outpatient clinic. She first developed symptoms in her early 20s with right lower limb dystonia and a rest-tremor involving her right upper limb. These symptoms had slowly progressed over the subsequent 30 years, with some general slowing of mobility and increased difficulty with fine motor tasks. Clinical examination found evidence of a rest-tremor, bradykinesia and generalized hyperreflexia. Within a few years of symptom onset, she was started on low-dose levodopa treatment (62.5 mg tds, with a good response, although dyskinetic movements had developed over the preceding 5–10 years. There were no reports of marked cognitive impairment nor mood change, and no preceding anosmia or symptoms of rapid eye movements (REM)-sleep behavioral disorder.*

*Parkin* (PARK2) is responsible for half of all cases of early-onset Parkinson's (under 45 years of age) and a significant proportion of sporadic cases. The typical clinical phenotype includes early-onset dystonia, hyperreflexia, and slower disease progression [5]. *Parkin* is believed to function as an E3-ligase, responsible for protein degradation, as well as mediating mitophagy (autophagic removal of damaged organelles within the mitochondria), likely involved in Parkinson's pathogenesis. A variety of mutation types have been identified to date, including missense mutations (most common form) and copy number variants (deletions and duplications).

#### 27.2.3.2 *PINK1/PARK6: PTEN-induced kinase 1*

*PINK1* mutations have been identified in 1%–2% of early-onset disease cases. The *PINK1* gene encodes a PTEN-induced kinase protein containing a protein kinase domain and a mitochondrial-targeting motif [6]. *PINK1* shares a number of overlapping functions and mechanistic pathways with other known Parkinsonian genes, principally *Parkin* and *DJ1*. In physiological form, *PINK1* accumulates on the outer membrane of dysfunctional mitochondria and its kinase activity is required for *Parkin* to be translocated into damaged mitochondria to facilitate mitophagy [7].

#### 27.2.3.3 *DJ1/PARK7: protein deglycase*

*DJ1* mutations are a rare cause of Parkinson's disease, accounting for <1% of early-onset parkinsonism, with missense mutations and deletions being most common [8,9]. Located on chromosome 1p36, *DJ1* is a redox sensor of oxidative stress and is likely to play a role in protecting neurons from oxidative stress-induced damage [10,11].

#### 27.2.3.4 *Glucocerebrosidase mutations*

Although normally associated with Gaucher's disease (a lysosomal storage disorder) in its homozygous form, heterozygous glucocerebrosidase (*GBA*) mutations are a strong risk factor for Parkinson's. Located on chromosome 1p22, *GBA* encodes a lysosomal enzyme that hydrolyzes glucocerebroside in glycolipid metabolism. It is thought that glucosidase activity helps modulate alpha-synuclein processing, with mutations causing accumulation of glycolipid metabolism intermediates, which then mediate cell damage [10]. *GBA* mutations are twice as common in early-onset versus late-onset Parkinson's and are associated with higher prevalence of cognitive impairment [10,12].

### 27.2.3.5 Other genes implicated in Parkinson's

Several other autosomal dominantly inherited genes have also been implicated in Parkinson's pathogenesis. *DNAJC13* encodes a chaperone protein that is thought to regulate protein trafficking and has mostly been found in those of Dutch-German-Russian Mennonite ancestry [13]. *CHCHD2* encodes a mitochondrial protein and has been found in a small cohort of Japanese families. Finally, links with *EIF4G1* have also been identified, although little is known of its function.

### 27.2.4 Treatment of Parkinson's disease

Treatment of both the motor and nonmotor symptoms in Parkinson's is symptomatic, with the overall emphasis being to increase the levels of midbrain dopamine. The two main groups of treatments for the management of motor symptoms include levodopa, also used in combination with a DOPA-decarboxylase inhibitor, and dopamine agonists, which target striatal dopaminergic neurons [13]. Other medications include monoamine oxidase type B inhibitors (e.g., selegiline), catechol-*O*-methyltransferase inhibitors (e.g., entacapone), amantadine, and anticholinergics [13]. In addition to oral medical therapy, treatment for Parkinson's can also be given via continuous infusion (e.g., apomorphine) or further down the gastrointestinal tract (e.g., jejunal duodopa). In eligible patients with refractory disease or the development of drug-induced dyskinesias, deep brain stimulation (DBS) provides a surgical means of symptoms management [14,15].

### 27.2.5 Parkinson's disease: key learning points

- There is a broad clinical phenotypic spectrum observed in genetically determined forms of Parkinson's disease.
- Autosomal-dominant inheritance of *LRRK2* mutations represent the most common form of genetically inherited Parkinson's disease.
- Common mechanistic pathways have been identified in *Parkin*, *PINK1*, and *DJ1* mutations.
- Heterozygous GBA mutations represent a risk allele for the development of parkinsonian symptoms but are not currently recognized as being directly responsible in giving rise to the disease phenotype.

## 27.3 Dystonia

### 27.3.1 Clinical characteristics

Dystonia is a movement disorder characterized by sustained or intermittent muscle contractions causing abnormal, often repetitive, movements, postures or both [16]. There is widespread variation in the clinical presentation of dystonia, with childhood, adolescent, and adult-onset forms, and muscle involvement ranging from single isolated groups (e.g., cervical dystonia) to more generalized forms. Many forms of dystonia have associated features, encompassing other movement disorders (such as parkinsonism), other neurological symptoms (such as peripheral neuropathies), and nonmotor associations. The association of psychiatric conditions is a specific common phenotype across dystonia disorders, for example, individuals with myoclonus dystonia (DYT11) demonstrate higher rates of generalized anxiety disorder, obsessive compulsive disorder, social phobia, and alcohol dependence compared to controls [17].

The most recent classification system aims to take into account diagnosis, treatment, underlying pathophysiology, and mechanism, organizing the dystonias according to two main features: clinical characteristics and etiology:

- Clinical features includes age at onset, regions of the body affected, temporal pattern of symptoms, presence of other movement disorders and other associated neurological features.
- Etiological categorization considers pathological nervous system findings (degeneration, structural nondegenerative lesions or neither) and inheritance pattern.

Dystonia is also divided into inherited, acquired, and idiopathic types. Acquired types are those noninherited forms with a known cause, such as dystonic cerebral palsy caused by brain injury in the perinatal period, and medication-induced forms. The idiopathic dystonias are those with no identified cause but may represent sporadic or currently unidentified genetic causes. The third category, which will form the focus of this section, are



the genetic dystonias. These are subdivided by mode of inheritance into autosomal dominant, autosomal recessive, X-linked recessive, and mitochondrial forms. Within the inherited dystonias, the DYT classification of diseases according to the locus of gene mutation can be helpful for subclassifying types.

### 27.3.2 Genetics of dystonia

With the use of next-generation genetic sequencing techniques, there has been a significant increase in the number of disease-causing genes identified for dystonia, with over 20 now recognized as having a role in disease pathogenesis. However, even within genetically homogeneous groups, there is substantial intra- and interfamilial variability in clinical phenotype and response to treatment. A number of genes linked with dystonia also demonstrate pleiotropic effects. An example of this is the gene implicated in paroxysmal kinesigenic dyskinesia, *PRRT2*. Encoding the proline-rich transmembrane protein 2, mutations in this gene locus are responsible for benign familial infantile epilepsy, infantile convulsions with choreoathetosis, and hemiplegic migraine. Even individuals within the same family group can exhibit phenotypical heterogeneity, with multiple different clinical conditions [18].

Dystonia genes demonstrate autosomal dominant, autosomal recessive, and X-linked inheritance patterns, with several of the dominantly inherited genes also displaying incomplete penetrance, such as *TorsinA* (DYT1), *THAP1* (DYT6), and *SGCE* (DYT11). The cause for this variation in penetrance of DYT1 (30%) and DYT6 (60%) remains largely unknown, although in DYT1 mutation carriers evidence is the emergence of additional genetic factors that may be of influence. A polymorphism in the cDNA at nucleotide 646, with a guanosine to cytosine substitution, encoding histidine rather than aspartate has been identified in 15% of cohorts. While no functional consequences directly relating to this have been identified, individuals carrying both mutation types have a higher likelihood of developing dystonia than those with the DYT1 mutation alone. *SGCE* (DYT11) also demonstrates variable penetrance via maternal imprinting mechanisms. Here, maternally inherited pathogenic mutations are silenced due to imprinting mechanisms, although may be subsequently expressed in later generations. In contrast, paternally inherited *SGCE* mutations are fully penetrant in demonstrating the disease phenotype [19].

### 27.3.3 Pathophysiology of dystonia

The pathophysiology of dystonia remains largely unknown; however, dystonia is widely considered to be a circuit-based disorder with disruption at the synaptic level impacting pathways, circuits, and networks. Evidence to date suggests that two circuits may be of particular importance, the basal ganglia-thalamo-cortical and cerebello-thalamo-cortical pathways, with imaging studies suggesting that disruption to these pathways may be critical in giving rise to dystonia [20]. The functioning of these connections are of particular importance, with diffusion tensor imaging demonstrating disrupted connections between motor cortical regions and subcortical connecting areas in those with focal dystonia. In addition, animal models have demonstrated reversible induced movement disorders including dystonia with pharmacological lesions of the basal ganglia [21]. On a cellular level, disruption to neurotransmitter function is also likely to contribute, particularly with the number of dystonia genes with direct links to the dopamine synthetic pathway, for example, dopa-responsive dystonia (DYT5, *GCH1* mutations). There is also accumulating evidence for the importance of cholinergic interneurons from animal models of dystonia, and disruption to serotonin metabolic pathways in human studies [22].

### 27.3.4 Targeted molecular diagnosis and therapy of dystonia

The clinical and genetic characteristics of the genetically determined forms of dystonia are summarized in Table 27.2.

### 27.3.5 Diagnosis of dystonia

The diagnosis of dystonia remains clinical; however, genetic testing can be used to provide greater specificity of diagnosis, further aiding management and genetic counseling. The majority of dystonia genes are now available for testing as gene panels, facilitating simultaneous testing of multiple potential causative genes and reducing the waiting time for patients to receive a genetic outcome [23].

**TABLE 27.2** Clinical and genetic description of dystonic disorders where the disease-causing gene has been identified.

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype	
						Motor	Nonmotor
DYT1 (605204)	0.17/100,000 mutation frequency	AD, 30% penetrance	TOR1A (9q32- q34)	Torsin1A	Torsin1A loss of function and increased degradation	Varies from focal to generalized dystonia. Usually early onset	Increased rate recurrent major depression
DYT2 (224500)	Unknown	AR	Likely HPAC (1p35.1)	Unknown	altered regulation of voltage- dependent calcium channels	Childhood or adolescent onset initially distal limb dystonia, slowly progressive	
DYT3 XDP/Lubag's disease (313650)	High rates in Philippine populations	X-linked	TAF1 (X13.1)	TAF1	Impaired cellular transcription- progressive neostriatal neuronal loss	Torsion dystonia. Progressive, severe. parkinsonian features	Increased rates of depression and anxiety
DYT4 Whispering dysphonia (602662)	Unknown	AD	TUBB4A (19p)	Tubb4a	Brain specific interference with assembly of tubulin subunits	Second to third decade progressive laryngeal dysphonia followed by generalization	Some alcohol and propranolol responsiveness Thin face and body habitus
DYT 5a Segawa's disease (600225)	0.5/1,000,000 More females than males	AD	GCH1 (14q22.1–22.2)	GTP cyclohydrolase 1	Reduced dopamine levels in substantia nigra	Young onset progressive limb dystonia. Improved by sleep, worse in evenings. Associated parkinsonism	Levodopa responsiveness. Neuropsychiatric features including depression, anxiety and OCD
DYT 5b (605407)	rare	AR	TH (11p15.5)	Tyrosine hydroxylase	Altered regulation of dopamine production	Two phenotypes: (1) perinatal severe encephalopathy with diurnal variation, (2) First year onset progressive generalized dystonia, hypokinesia, rigidity	Autonomic disturbances
DYT 5b (612716)	Rare	AR	SPR (2p13.2)	Sepiapterin reductase	Defect in BH4 synthesis causing severe deficiencies in dopamine and serotonin	Usually starts in lower limbs. Diurnal fluctuation, levodopa responsive. Parkinsonism. Ataxia	Cognitive delay. Psychiatric and behavioral abnormalities
DYT6 (609520)	Unknown	AD, ~60% penetrance	THAP1 (8p11.1)	THAP domain encoding protein	Involved in endothelial cell proliferation and proapoptotic processes	Early craniofacial involvement with secondary generalization. Laryngeal dystonia common	
DYT8 Paroxysmal non- kinesigenic Dyskinesia (609023)	Rare	AD	MR1 (2q35)	MR-1L	?altered glutathione related detoxification in neuronal cells	Onset infancy or childhood. Episodic dystonia, choreoathetosis, ballism. Precipitated by stress, caffeine, alcohol, exertion, ovulation, and menstruation	

DYT10 Paroxysmal kinesigenic dyskinesia (614386)	1/150,000	AD, incomplete penetrance	PRRT2 (16p11.2- q12.1)	Proline-rich transmembrane protein 2	Protein highly expressed in brain and spinal cord, interacts with a synaptosomal membrane	Childhood onset, brief attacks of dystonia, chorea, and athetosis. Improves with age and carbamazepine	Depressive and behavioral disorders
DYT 11 Myoclonus dystonia (604149)	Rare	AD, incomplete penetrance- maternal imprinting	SGCE (7q21.3)	Epsilon-sarcoglycan	Reduced expression at the cell surface membrane. Potential role in synaptic plasticity	Childhood/Adolescent onset, dystonia affects mainly neck, trunk and arms. Myoclonic jerks. Improved by ethanol and clonazepam	Neuropsychiatric disorders including depression, OCD, personality disorder
DYT12 Dystonia- parkinsonism (182350)	Rare	AD	ATP1A3 (19q13.2)	Alpha 3 subunit of Na/K ATPase	Impaired activity of enzyme involved in regulation of neuronal activity, expressed in regions including basal ganglia, hippocampus, and cerebellum	Young adults. Sudden onset, asymmetrical dystonia and parkinsonism	Neuropsychiatric disorders including psychosis, anxiety, and depression
DYT16 (603424)	Reported in two Brazilian families	AR	PRKRA (2q31.2)	Interferon-inducible double-stranded RNA-dependent protein kinase activator A	?disrupted cellular stress response	Early-onset dystonia and parkinsonism	Aggression. Cognitive impairment
DYT18 GLUT1 Deficiency Syndrome 2 (138140)	1 in 90,000	AD	SLC2A1 (1p34.2)	GLUT1	Impairment of glucose transport into the brain	Episodic dyskinesia, mainly distal lower limb dystonia/choreoathetosis. Triggers include hunger, exercise. Can get ataxia, spasticity, seizures, encephalopathy	Associated with hemolytic anemia. Improves with a ketotic diet
DYT23 (614860)	Reported in one German family and one Dutch family	AD	CACNA1B (9q34.3)	CACNA1B	Alteration in neurotransmitter release at synapses	Adult-onset cervical dystonia, tremor of head and limbs one family. Myoclonic dystonia the other family	
DYT24 (615034)	Unknown	AD	ANO3 (11p)	Anoctmin-3	A transmembrane calcium activated protein channel, highly expressed in putamen. Reduced intracellular calcium signaling	Adult-onset cervical dystonia, upper limb tremor, and laryngeal involvement	
DYT25 (615073)	Unknown	AD	GNAL (18p)	G-protein G alpha subunit	Highly expressed in basal ganglia, mutations related to abnormality in dopamine D1 receptor activity	Adult onset, focal cranial and cervical dystonia	
DYT26 (616398)	Unknown	AD	KCTD17 (22q12.3)	Potassium channel tetramerization- domain containing proteins	Highly expressed in putamen. ?role in dopaminergic dysfunction and cellular calcium homeostasis	Onset of mild upper limb dystonia, progressive. Myoclonic jerks	
DYT28 (617284)	Unknown	AD	KMT2B (19p13), reduced penetrance	Histone-lysine N methyltransferase 2B	Involved in regulation of transcription	Childhood onset, progressive	Elongated face, bulbous nose. Developmental delay and intellectual disability

(Continued)

**TABLE 27.2** (Continued)

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype	
						Motor	Nonmotor
Aromatic-L- amino acid decarboxylase deficiency (608643)	Rare	AR	DDC (7p12)	Aromatic L-amino acid decarboxylase	Deficiency of protein that catalyzes dopamine synthesis	Onset in infancy or childhood of hypotonia, dystonia, oculogyric crises, developmental delay	Irritability and emotional lability
Dopamine transporter deficiency syndrome (613135)	Rare	AR	SLC6A3 (5p15.33)	Sodium-dependent dopamine transporter	Defective presynaptic dopamine reuptake	Onset in infancy or childhood of parkinsonism and dystonia	Irritability

*AD*, Autosomal dominant; *AR*, autosomal recessive; *OCD*, obsessive compulsive disorder.

### 27.3.5.1 DYT1: TorsinA mutations

#### Case report

An 8-year old girl was referred to the pediatric neurology out-patient clinic by her general practitioner, attending with her parents. She was born at term, at the end of an uneventful pregnancy. There was no report of perinatal injury or anoxic insult, and no developmental delay in infancy or early childhood. Two years prior to presentation she had begun to develop pain and cramping in her right foot, progressing to abnormal posturing, particularly inversion and rotation of the foot. During this time, walking became more difficult, with a tendency to walk on the outer aspect of her foot. Over the preceding 6 months similar symptoms had begun to develop in the left lower limb and to a lesser extent the right upper limb. There was no immediate family history of similar symptoms, although her paternal grandfather had had lifelong mobility difficulties, similar posturing of the limbs and symptoms of major depression in adult life. Cerebral imaging and routine blood tests (including copper and caeruloplasmin) were within normal limits, with no evidence of focal pathology.

DYT1 dystonia is caused by mutations to the *TOR1A* gene, resulting in loss of function of the encoded protein, Torsin1A. The most common mutation type is a three base pair deletion (GAG) at the c.907\_909 locus, with this accounting for 16%–53% of all types of early-onset dystonia in non-Jewish populations, and 80%–90% in Ashkenazi Jewish populations. No genotype–phenotype correlation has been identified to date, although the link with the guanosine to cytosine genetic variant described earlier provides some evidence of the genetic underpinning of the observed variable penetrance. Murine and *Drosophila* models of this disorder suggest that *TOR1A* mutations result in a loss of function of the Torsin1A protein, primarily due to destabilization and premature degradation, resulting in failure of its expression. In addition, although no focal areas of pathology have been identified, animal models carrying this mutation have demonstrated some evidence of abnormal nuclear membranes in postmitotic neurons.

### 27.3.5.2 DYT5: GCH1 mutations (dopa-responsive dystonia)

#### Case report

An 11-year old boy was referred to the pediatric neurology out-patient department, attending with both of his parents. He had been born at term, at the end of a normal pregnancy and without any perinatal complications. There were no reports of developmental delay in infancy, and no history of recurrent hospital admissions during this period. At the age of 6 years, he began to develop abnormal posturing of his right foot, with evidence of inversion and a tendency to walk on the outer aspect of his foot. This became more pronounced with running and was impacting his ability to partake in sporting activities at school. Over the past 5 years these symptoms had generalized to involve his remaining limbs and trunk, limiting all activities of daily living. These symptoms tended to be worse toward the end of the day and improved after sleep (diurnal variation). This boy had three siblings, one, his younger sister having developed a similar spectrum of symptoms. Six months prior to this appointment he had been started on levodopa treatment with a dramatic improvement to his limb posturing and general mobility.

Dopa-responsive dystonia (DYT5) can be caused by mutations to the *GCH1*, *TH*, and *SPR* genes. Mutations in the *GCH1* gene (previously referred to as Segawa's disease) are the most common and are inherited in an autosomal-dominant fashion. *GCH1* encodes the GTP cyclohydrolase 1 enzyme, critical in the catalysis of the first step in the synthesis of tetrahydrobiopterin (TH4). Tetrahydrobiopterin itself is an essential cofactor in amino acid processing and neurotransmitter synthesis. This latter role includes monoaminergic neurotransmitters, and in particular dopamine. Pathological studies have demonstrated a reduction of dopamine levels in the substantia nigra of patients with DYT5 dopa-responsive dystonia in the absence of any neurodegeneration, and normal levels of tyrosine hydroxylase activity.

### 27.3.5.3 DYT6: THAP1 mutations

#### Case report

A 20-year old man was seen in the movement disorders outpatient clinic with a 2-year history of spasms and unusual movements involving the face, and in particular his eyes. He described these as being involuntary with no preceding warning and would frequently cause the muscles around his eyes and mouth to contract and “spasm.” The symptoms involving the eyes were particularly pronounced with excess blinking and periods of sustained eye closure. These symptoms had progressed over the past 12 months to involve his speech and swallow. The phonation of his speech had become more variable and on occasion sounded as though he was breathless, while he had also begun to choke when drinking liquids. In the 6 months leading up to the appointment he had also noted cramping in his right hand and forearm when writing.

THAP1 mutations are inherited in autosomal-dominant fashion but with only ~60% penetrance, the cause for this being unknown. THAP1 itself encode the THAP1 protein that functions as a DNA-binding transcriptional



regulator that regulates endothelial cell proliferation and G1/S-cell cycle progression. It is also thought that it may have some proapoptotic activity, aiding cell apoptosis and degradation. Functional cellular studies have suggested that *THAP1* mutations impair protein stability, reducing the amount of function THAP1 protein that is available to bind to DNA, and therefore reducing its cellular function.

#### **27.3.5.4 DYT10: paroxysmal kinesigenic dyskinesia**

##### *Case report*

A 10-year old boy attended the neurology out-patient clinic with his parents. Over the preceding 6 months he had begun to develop “attacks” of abnormal movements. These tended to occur after a sudden movement, for example, standing up quickly from a chair, or suddenly starting to run in the school playground. The attacks involved abnormal posturing of his lower limbs bilaterally, with a tendency of his feet to invert. He also described cramping and pain in his lower limbs during these events. Each attack would last 30–40 s, and could occur up to 40–50 times per day dependent upon the activities undertaken. There was no reported loss of consciousness during the events. The patient’s father also described similar, less frequent attacks, involving his upper limbs the description of which was more in keeping with chorea. On examination there was no evidence of a focal neurological deficit in either the patient or his father. The young boy had been prescribed Carbamazepine by his general practitioner, with a significant improvement to his symptoms.

Paroxysmal kinesigenic dyskinesia is caused by mutations to the proline-rich transmembrane protein 2 (PRRT2), which are inherited in an autosomal-dominant manner. PRRT2 mutations demonstrate pleiotropy, giving rise to a number of distinct clinical disorders including infantile convulsions and choreoathetosis Syndrome, benign familial infantile convulsions, and hemiplegic migraine. The proline-rich transmembrane protein 2 itself is predicted to include two transmembrane segments. Its function in the brain remains unknown; however, it is believed to interact with the synaptosomal-associated protein 25 (SNAP25), with a suggestion that it may play a role in the fusion of synaptic vesicles to the plasma membrane. Patients with PRRT2 mutation positive paroxysmal kinesigenic dyskinesia often respond well to treatment with carbamazepine and phenytoin, resulting in a significant reduction to the number of events experienced.

#### **27.3.5.5 DYT11: SGCE mutations**

##### *Case report*

A 7-year old girl is reported to have developed jerks involving her upper body at the age of 3 years. These were most evident when feeding or drawing. There was no evidence that these were stimulus sensitive and were not associated with a loss of consciousness or awareness. The jerks had remained relatively stable over the preceding 4 years, with no spread to other body regions. More latterly she had also begun to develop some abnormal posturing of her neck and upper limbs, again more pronounced when undertaking tasks. Her parents described some anxiety-related symptoms, particularly in relation to unfamiliar social settings. Both of her parents were well with no evidence of a movement disorder. However, her maternal grandmother had developed similar symptoms in childhood that had persisted into adult life. Her motor symptoms were highly alcohol responsive, almost completely resolving after a few glasses of wine. In addition, she also had diagnoses of obsessive-compulsive disorder and generalized anxiety disorder and was reviewed regularly by the local psychiatry department.

Myoclonus dystonia is caused by mutations to the epsilon-sarcoglycan (SGCE) gene located on chromosome 7 and encoding the epsilon-sarcoglycan protein, a single-pass transmembrane protein. The role of the epsilon-sarcoglycan protein in brain remains unknown, but in peripheral tissue forms part of the dystrophin-associated glycoprotein complex. Hypothetical models have suggested that the epsilon-sarcoglycan protein may be preferentially located on the postsynaptic membrane and play a role in postsynaptic receptor clustering or function. Mutations in the SGCE gene result in degradation of the protein by the ubiquitin-proteasome system, and failure of its expression at the cell surface membrane. Postmortem tissue analysis has demonstrated that the highest levels of epsilon-sarcoglycan protein are expressed in the basal ganglia and cerebellum, with evidence from cerebral imaging demonstrating a positive correlation between dystonia severity and gray matter volume, as well as reduced striatal dopamine-2-receptor binding in those with SGCE mutations.

### **27.3.6 Therapy of dystonia**

The currently available treatment for dystonia involves physical therapy, oral medication, localized neurotoxin injections, and DBS.

### 27.3.6.1 Physiotherapy

Physiotherapy is used to aid symptomatic management in dystonia, helping with control of pain, improvement of posturing of the affected area, and aiding of day-to-day functioning. Techniques used include passive mobilization and stretching of affected areas, and training aimed at improving voluntary control and functionally relevant activities.

### 27.3.6.2 Oral medication

Medical treatment of dystonia focuses on the neurotransmitter systems implicated in dystonia pathophysiology. These systemic forms of treatment are predominantly used in the more generalized dystonia phenotypes rather than for focal dystonias where more targeted treatments are usually utilized. Anticholinergic medications such as trihexyphenidyl, with their action as postsynaptic antagonists of muscarinic receptors, provide symptomatic treatment in some, although potential side effects include a dry mouth, urinary retention, constipation, impairment in concentration and memory, and confusion. Benzodiazepines impact gamma aminobutyric acid (GABA) neurotransmission by binding to GABA<sub>A</sub> receptors, increasing the opening of chloride channels and resulting in enhancement of inhibitory neurotransmission. Clonazepam is a commonly used benzodiazepine in dystonia management due to its long half-life. Side effects include sedation, disinhibited behavior, depression, and drooling. In the specific case of dopa-responsive dystonia (DYT5, *GCH1* mutations) a large and sustained improvement can be seen with administration of levodopa. Dopamine reducing medications also improve symptoms in dystonia. Examples of these include tetrabenazine, which reduces dopamine presynaptically; and clozapine that acts on the postsynaptic membrane to block dopamine uptake.

### 27.3.6.3 Botulinum toxin

Botulinum toxin therapy forms the first-line treatment for those with focal dystonia, in particular cervical dystonia. The botulinum toxin is injected into the affected area, exerting its effects locally by preventing presynaptic acetylcholine release and so relaxing the injected muscle. The effects of this treatment gradually wane, typically over a 12-week period, requiring repeated treatment for sustained effect.

### 27.3.6.4 Deep brain stimulation

DBS is indicated in those with generalized or segmental dystonia and complex cervical dystonia, where botulinum toxin treatment is ineffective or its effect has waned. It is a surgical procedure involving the insertion of electrodes, which can apply stimulation in functionally relevant brain areas. The globus pallidus interna, a subsection of the basal ganglia, represents the most commonly targeted area for stimulation, while evidence also supports the stimulation of the thalamus and subthalamic nucleus. Using MRI guidance, the electrodes are placed bilaterally in the target location, together with an implantable pulse generator, placed below the left clavicle, acting as a pacemaker for the electrical stimulation. The intensity of stimulation can gradually be increased to reach a level that optimizes benefit but avoids side effects. It takes several months before the beneficial effects are experienced. Risks of the procedure are similar to those inherent in all neurosurgical procedures, including bleeding, infection, and adverse effects on other functional brain regions [24].

## 27.3.7 Dystonia: key learning points

- Dystonia is often described according to its clinical distribution, including generalized, segmental, and focal forms.
- Proposed pathophysiological mechanisms include disrupted synaptic neurotransmission impacting neuronal circuits and networks.
- Genetically determined forms of dystonia can be inherited in autosomal dominant, autosomal recessive, X-linked recess, and via mitochondrial mutations
- Several of those with autosomal-dominant inheritance demonstrate reduced penetrance (e.g., DYT1, DYT6, and DYT11).
- Within dystonia genetics, there are some examples of pleiotropy, that is, mutations to the same gene result in distinct clinical disorders, for example, *PRRT2* mutations.

## 27.4 Ataxia

Ataxia describes a loss of coordination of movement, which can affect the limbs and trunk [25]. There are many heritable forms of ataxia, where in addition to the loss of coordination additional clinical characteristics may also be present. A summary of the clinical, genetic and proteomic characteristics of genetically determined forms of ataxia can be seen in Table 27.3 (spinocerebellar ataxias) and Table 27.4 (other forms of ataxia).

**TABLE 27.3** Clinical and genetic description of the spinocerebellar ataxias (SCA).

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype
SCA1 (164400)	rare	AD	<i>ATXN1</i> (6p22) (CAG repeat expansion)	Ataxin-1	Disrupted transcriptional regulation	Pyramidal signs, peripheral neuropathy, chorea
SCA2 (183090)	Unknown. Cuba Holguin Province 43/100,000	AD	<i>ATXN2</i> (12q24.12) (CAG repeat expansion)	Ataxin-2	Role of ataxin-2 in RNA metabolism	Slowed saccades, parkinsonism, and dementia
SCA3 Machado–Joseph disease (109150)	Commonest SCA	AD	<i>ATXN3</i> (14q32.12) (CAG repeat expansion)	Ataxin-3	Protein misfolding, accumulation of disease protein in neuron inclusions	Cerebellar ataxia, pyramidal and extrapyramidal signs, fasciculations, oculomotor apraxia
SCA5 (600224)	Rare	AD	<i>SPTBN2</i> (11q13.2)	Beta-III spectrin	Calcium-mediated alteration of transcription	Slowly progressive. Eye movement abnormalities, tremor, impaired vibration sense
SCA6 (183086)	<1/100,000	AD	<i>CACNA1A</i> (CAG repeat expansion)	Alpha 1a subunit voltage-gated calcium channel	Disrupted calcium channel function	Predominantly cerebellar symptoms. Slow progression
SCA7 (164500)	Rare	AD	<i>ATXN7</i> (CAG repeat expansion)	Ataxin-7	Impaired transcription. Development of neuronal nuclear inclusions	Visual loss (cone-rod dystrophy), slow saccades
SCA8 (608768)	Unknown	AD	<i>ATXN8/ATXN80S</i> (13q21) (CTG repeat expansion)	Ataxin-8	RNA-related neurotoxicity	Cognitive dysfunction, psychiatric disorders. Pyramidal and sensory signs
SCA10 (603516)	unknown	AD	<i>ATXN10</i> (22q13.31) (ATTCT pentanucleotide expansion)	Ataxin-10	RNA processing abnormality	Slowly progressive cerebellar symptoms and epilepsy. Mild pyramidal signs, peripheral neuropathy
SCA11 (604432)	Rare	AD	<i>TTBK2</i> (15q15.2)	Tau tubulin kinase	Reduction of tau protein phosphorylation, causing tau deposition	Slowly progressive cerebellar symptoms and eye movement abnormalities. Occasional hyperreflexia, dystonia and parkinsonism
SCA12 (604326)	Unknown/Rare	AD	<i>PPP2R2B</i> (5q32) (CAG repeat expansion)	Serine/Threonine-protein phosphatase		Mild ataxia. Action tremor. Pyramidal and extrapyramidal signs. Dementia

(Continued)

TABLE 27.3 (Continued)

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype
SCA13 (604259)	Unknown	AD	<i>KCNC3</i> (19q13.33)	Voltage gated potassium channel	Disrupted voltage gated potassium channel function. Altered neuronal excitability	Childhood onset: developmental delay, mild ataxia. Short stature. Late development: dysphagia, bradykinesia, and urinary urgency
SCA14 (605361)	Rare	AD	<i>PRKCG</i> (19q13.42)	Protein kinase C $\gamma$	Aggregation of abnormal protein in Purkinje cells	Variable age at onset hyperreflexia, reduced vibration sensation
SCA15 (606658)	Unknown	AD	<i>ITPR1</i> (3p26.1)	IP3 receptor	Disruption to calcium channel function impacting synaptic signaling	Slowly progressive. Head tremor
SCA17 (607136)	Unknown	AD	<i>TBP</i> (6q27) (CAG repeat expansion)	TATA binding protein	Impaired transcriptional regulation.	Dementia, parkinsonism, dystonia, epilepsy, chorea, spasticity, and psychiatric disorders
SCA19/22 (607346)	Unknown	AD	<i>KCND3</i> (1p13.2)	Potassium channel	Voltage gated potassium channel dysfunction	Myoclonus, postural tremor, cognitive impairment
SCA21 (607454)	Unknown	AD	<i>TMEM240</i> (1p36.33)	Transmembrane protein 240	Unknown	Parkinsonism, mild cognitive impairment
SCA23 (610245)	Unknown	AD	<i>PDYN</i> (20p13)	Proenkephalin-B	Increased NMDA receptor signaling leading to cellular dysfunction and death	Slowing of saccades, ocular dysmetria, dysarthria, and hyperreflexia
SCA26 (609306)	Reported in one American family	AD	<i>EEF2</i> (19p13.3)	Elongation factor 2	Impaired translocation	Slowly progressive. Nystagmus and impaired pursuit
SCA27 (609307)	Rare	AD	<i>FGF14</i> (13q33.1)	Fibroblast growth factor 14	Abnormality of presynaptic calcium channel regulation	Tremor, dyskinesia
SCA28 (610246)	Rare	AD	<i>AFG3L2</i> (18p11.22)	AFG3-like protein 2	Purkinje degeneration	Juvenile onset, slow progression. Later pyramidal signs, ptosis, slowing of saccades, ophthalmoparesis
SCA29 (117360)	Unknown	AD	<i>ITPR1</i> (3p26.1)	Inositol 1,4,5-triphosphate receptor type 1	Abnormalities of membrane channel function	Onset at birth, slow or no progression
SCA31 (117210)	Rare	AD (incomplete penetrance)	<i>BEAN1</i> (16q21)	Protein BEAN1	RNA abnormalities	Eye movement abnormalities
SCA35 (613908)	Reported in 3 Chinese families	AD	<i>TGM6</i> (20p13)	Transglutaminase 6	Reduction of intranuclear TG6, accumulation in perinuclear region	Hyperreflexia, extensor plantars, and spasmodic torticollis
SCA36 (614153)	Unknown	AD	<i>NOP56</i> (20p13)	Nucleolar protein 56	RNA function abnormality, toxic effect	Hearing loss, tongue atrophy, fasciculations, and peripheral neuropathy

AD, Autosomal dominant; AR, autosomal recessive.

**TABLE 27.4** Clinical and genetic description of other forms of inherited ataxia.

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype
Friedreich ataxia (229300)	Commonest autosomal recessive ataxia	AR	<i>FXN</i> (9q21.11) (GAA repeat expansion)	Frataxin, mitochondrial	Mitochondrial iron accumulation leading to oxidative stress.	Limb weakness, absent lower limb reflexes, upgoing plantars. Dysarthria, pes cavus, scoliosis, visual impairment, and cardiomyopathy
DRPLA (125370)	Most common in Japan	AD	<i>ATN1</i> (12p13.31) (CAG repeat expansion)	Atrophin 1	Reduced transcription of fat tumor suppressor gene, leading to neuronal degeneration	Onset <20 years: myoclonic epilepsy and intellectual disability Onset >40 years: choreoathetosis, dementia
EA1 (160120)	unknown	AD	<i>KCNA1</i> (12p13.32)	Potassium voltage-gated channel subfamily A member 1	Reduction of potassium channel function leading to increased neuronal excitability.	Brief episodes of ataxia. Myokymia, episodic muscle contractions
EA2 (108500)	Unknown	AD	<i>CACNA1A</i> (19p13.13)	Voltage-dependent P/Q-type calcium channel subunit alpha 1A	Abnormal function of calcium channel heavily expressed in the cerebellum	Episodic ataxia. Nystagmus, vertigo, migraines
EA5 (613855)	Single French-Canadian family	AD (incomplete penetrance)	<i>CACNB4</i> (2q23.3)	Voltage-dependent L-type calcium channel subunit beta-4	Calcium channel dysfunction	Episodic vertigo and ataxia. Nystagmus between episodes
EA6 (612656)	Three unrelated families identified	AD	<i>SLC1A3</i> (5p13.2)	Excitatory amino acid transporter 1	Reduced glutamate transport	Episodic and progressive ataxia. Hemiplegia, seizures
Cerebellar ataxia, deafness and narcolepsy (604121)	Reported in four families	AD	<i>DNMT1</i> (19q13.2)	DNA (cytosine-5)-methyltransferase 1	Abnormal DNA methylation	Deafness, narcolepsy, dementia. Sometimes optic atrophy, sensory neuropathy, depression, psychosis
Ataxia-telangiectasia (208900)	1/40,000–300,000	AR	<i>ATM</i> (11q22.3)	Serine-protein kinase ATM	Lack of repair of damaged DNA and increased oxidative stress	Childhood onset, oculomotor apraxia, cutaneous telangiectasia, immune deficiency, hypogonadism, and hematological malignancies
Ataxia with vitamin E deficiency (277460)	Unknown	AR	<i>TTPA</i> (8q12.3)	Alpha-tocopherol transfer protein	Low levels of vitamin E, increased oxidative stress	Childhood onset, progressive. Dysarthria, peripheral neuropathy, loss of lower limb reflexes, retinitis pigmentosa
Ataxia with oculomotor apraxia type 1 (208920)	Unknown	AR	<i>APTX</i> (9p21.1)	Aprataxin	Impaired DNA repair	Childhood onset. Oculomotor ataxia, axonal peripheral neuropathy, hypoalbuminaemia
Ataxia with oculomotor apraxia type 2 (606002)	Unknown	AR	<i>SETX</i> (19q34.13)	Probable helicase senataxin	Implicated in DNA repair	Choreoathetosis, dystonic posturing, and oculomotor ataxia
Cerebrotendinous xanthomatosis (213700)	Rare	AR	<i>CYP27A1</i> (2q35)	Sterol 26-hydroxylase, mitochondrial	Defective bile acid biosynthesis	Onset early childhood. Spinal cord involvement, dementia, tendon xanthomas, juvenile cataracts, and early atherosclerosis

AD, Autosomal dominant; AR, autosomal recessive.



### 27.4.1 Genetics of ataxia

Hereditary ataxias demonstrate a range of inheritance patterns including autosomal dominant, autosomal recessive, X-linked, and inheritance of mitochondrial mutations. Within these groups, a large number constitute trinucleotide repeat disorders, where there is an expansion of a three nucleotide repeat, for example, CAG, to pathogenic levels. As with other trinucleotide repeat disorders, these disorders demonstrate genetic anticipation, in which the size of the repeat increases with successive generations with a progressively earlier age at onset.

#### 27.4.1.1 Gene transcription and RNA

Many of the proteins expressed by the mutated genes in the hereditary ataxias have a role in regulating gene expression (e.g., SCA1, SCA2, and SCA7) with a subsequent impact on neuronal cell function and ultimately degeneration. Trinucleotide repeat segments are commonly found within transcription regions, with several lines of evidence suggesting that these transcription regions impact histone acetylation, subsequently affecting chromatin regulation and gene expression [26]. One such example is SCA7, in which the ataxin protein typically forms part of the histone acetyltransferase complex, increasing acetylation of the histone H3, which in turn causes downregulation of transcription. Similarly, a polyglutamine expansion of the TATA box-binding protein in SCA17 results in protein dysfunction and subsequent disruption of the transcription initiation process [27]. Mutations in noncoding regions are also observed among the genetically determined ataxias. One such example is Friedreich's ataxia, caused by a GAA repeat expansion within an intronic region of the *frataxin* gene. Several mechanisms have been proposed for the resultant failure of transcription of this gene, including the physical obstruction of the DNA-polymerase II in unwinding DNA as well as that the abnormal expansion behaves like heterochromatin (chromatin too tightly bound to be translated effectively), which acts to silence the expression of neighboring regions.

#### 27.4.1.2 Intranuclear inclusions

A common neuropathological finding among the hereditary ataxias is neuronal accumulation of misfolded proteins, predominantly in the nucleus, and to a lesser extent in the cytoplasm of cells. Frequently these protein aggregates will include not only the pathological protein, but other proteomic elements such as transcription factors, which with the additional impairment of the ubiquitin-proteasome system, accumulate due to the failure of abnormal protein degradation. A more specific example of this is SCA3, where the ataxin-3 protein interacts with the ubiquitin-proteasome system, with accumulation of pathogenic ataxin-3 resulting in impairment to its normal function [28].

#### 27.4.1.3 Transmembrane channel abnormalities

Several types of ataxia involve mutations in genes encoding transmembrane channels, including voltage-gated calcium channels (SCA6); voltage-gated potassium channel (SCA13, EA1, and SCA19); ligand-gated calcium channels (SCA15); and active transporters (EA6). Overall, this disruption to channel function impacts neuronal excitability and subsequent neuronal activity [29].

#### 27.4.1.4 Neuronal calcium homeostasis

Synaptic neurotransmission requires maintenance of calcium homeostasis for normal function, with disruption to these processes implicated in the pathogenesis of some forms of hereditary ataxia. Calcium homeostasis abnormalities can be the direct result of calcium channel abnormalities (such as a mutation in the *IP3R1* smooth endoplasmic reticulum (ER) calcium channel in SCA15), or result indirectly from impaired calcium signaling, such as with SCA1. Calcium homeostasis is also important in intracellular signaling pathways, which in turn impact transcriptional regulation [29].

#### 27.4.1.5 Mitochondrial dysfunction

Many conditions caused by inheritance of mitochondrial DNA mutations involve ataxia as part of a broader phenotype, for example, Kearns–Sayres and myoclonic epilepsy with ragged red fibers. Several of the autosomal recessively inherited forms of ataxia also interfere with mitochondrial function, such as infantile onset spinocerebellar ataxia. Mitochondria function in the production of adenosine triphosphate (ATP), failure of which results in a deficit in cellular energy production with neurological sequelae including cerebellar dysfunction [29].

### 27.4.2 Cerebellar degeneration

The mechanisms discussed both directly and indirectly ultimately lead to cell death, with particular involvement of cerebellar Purkinje cell neurons via both apoptosis and necrotic mechanisms. Atrophy of the cerebellum, as well as other brain regions, is a notable feature of the hereditary ataxias, particularly the autosomal dominantly inherited forms. There is evidence that this selective degeneration may be related to cerebellar Purkinje cells being particularly vulnerable due to their high metabolic activity and the balance of protein production and degradation mechanisms [30] (Table 27.5).

### 27.4.3 Targeted molecular diagnosis and therapy

#### 27.4.3.1 Diagnostic testing

##### 27.4.3.1.1 Blood plasma tests

For some forms of ataxia a blood plasma test can be used to measure pathological levels of specific substrates (Table 27.6).

##### 27.4.3.1.2 Genetic testing

Molecular genetic testing is available for both the trinucleotide repeat disorders and the other forms of inherited ataxias. Due to the clinical overlap of many of the autosomal-dominant spinocerebellar ataxias, gene panel testing of several of the disorders simultaneously is frequently used. When a trinucleotide repeat disorder is confirmed, the length of the repeats is quantified to determine if this is within normal, intermediate, or pathogenic ranges. In some cases the size of the trinucleotide repeat may be too large to be determined using this technique, in which case Southern blotting would be used for further analysis.

### 27.4.4 Friedreich's ataxia

#### Case report

A 20-year old man attended the neurology outpatient clinic following a referral by his general practitioner. Five years ago he had begun to develop what was initially described as clumsiness, followed by increased difficulties with coordination, frequent falls and worsening of his mobility. Over this time, friends and family had noticed that his speech had become more slurred and that he was more difficult to understand. He also reported occasional episodes of breathlessness, although these were present only on exertion. On examination there was evidence of increased tone and sustained clonus in the lower limbs bilaterally. Examination of power demonstrated proximal limb weakness involving both the upper and lower limbs, deep tendon reflexes were absent throughout and plantar responses were upgoing bilaterally. There was also evidence of spinal scoliosis involving the lumbosacral region of the spine, and pes cavus deformities of the feet bilaterally. Cognitive examination identified no marked evidence of impairment. Echocardiogram and electrocardiogram recordings were both suggestive of hypertrophic cardiomyopathy. There was no reported family history of similar symptoms.

Friedreich's ataxia is caused by mutations to the *frataxin* (FXN) gene, inherited in an autosomal recessive manner. The *frataxin* gene encodes the frataxin protein, which is important in the normal functioning of mitochondria. Friedreich's ataxia is a trinucleotide repeat disorder, with expansion of the GAA repeat beyond the normal limits (normal range: 5–33 repeats). Typically >66 GAA repeats is seen in Friedreich's ataxia, with fewer repeats correlating with a later onset of motor symptoms. In the pathological setting, there is reduced production of the frataxin protein that leads to degeneration of sensory neurons at the dorsal root ganglion, with subsequent involvement of the spinal cord and cerebellum. Currently, there are no disease-modifying or curative therapies available for Friedreich's ataxia, with treatment principally being supportive and including walking aids, prostheses, and cardiac monitoring.

### 27.4.5 Spinocerebellar ataxia 2

#### Case report

A 30-year old man presented with a 2-year history of progressive loss of coordination, balance difficulties and falls. Approximately 12 months after onset of these symptoms he also began to develop difficulties with his speech and swallow, resulting in approximately a stone of weight loss. He also described sensory disturbance in his hands and feet, reporting both numbness and a "pins and needles" sensation, as well as a bilateral upper limb tremor. Three other family members were

**TABLE 27.5** Clinical and genetic description of other neurodegenerative movement disorders.

Disorder (MIM number)	Epidemiology	Inheritance	Causative gene (locus)	Protein product	Proposed impact of mutation	Clinical phenotype	
						Motor	Nonmotor
Huntington's disease 143100	Prevalence 5-7/100,000 in Caucasian population	AD	HTT (4p16.3)	Huntingtin	Medium spiny neuron damage	Chorea, dystonia	Cognitive impairment, psychosis
Wilson's disease 277900	30/1,000,000	AR	ATP7B (13q14.3)	Copper-transporting ATPase 2	Toxicity secondary to copper accumulation	Parkinsonism	Psychosis, depression, liver failure, hemolytic anemia
PKAN 234200	Rare	AR	PANK2 (20p13)	Protein Pantothenate kinase 2	?disordered CoA biosynthesis	Parkinsonism, dystonia, dysarthria, spasmodic dysphonia	Pigmentary retinopathy, mood lability, impulsivity, abnormal eye movements
PLAN 610217	Rare	AR	PLA2G6 (22q13.1)	Calcium-independent phospholipase A2	Uncertain	Dystonia, parkinsonism, spastic tetraparesis	Developmental regression, optic atrophy
Mitochondrial MPAN 614298	Rare	AR	C19orf12 (19q12)	Protein C19orf12	Possible dysfunctional mitochondrial autophagy and increased apoptosis	Dystonia of distal limbs, spasticity	Optic atrophy, cognitive impairment
BPAN 300894	Rare	X-linked dominant	WDR45 (Xp11.23)	WD repeat domain phosphoinositide-interacting protein 4	Possible impaired autophagy	Dystonia, parkinsonism	Developmental delay, seizures
FAHN 612319	Rare	AR	FAH2 (16q23.1)	Fatty acid 2-hydroxylase	Reduced 2-OH production	Gait disorder, spasticity	Frequent falls, cognitive decline, seizures, optic atrophy
CoPAN 615643	Rare	AR	COASY (17q21.2)	Bifunctional coenzyme A synthase	CoA deficiency	Dystonia, dysarthria	Cognitive decline
Neuroferritinopathy 606159	Rare	AD	FtL (19q13.33)	Ferritin light chain	Ferritin precipitation and aggregation	Action-induced and orofacial dystonia, parkinsonism, chorea	Cognitive decline
Aceruloplasminemia 604290	Rare	AR	CP (3q24-q25)	Ceruloplasmin	Iron toxicity	Facial dystonia	Diabetes, macrocytic anemia, cognitive decline
Woodhouse–Sakati syndrome 241080	Rare	AR	DCAF17 (2q31.1)	DDB1- and CUL4-associated factor 17	Uncertain	Parkinsonism	Hypogonadism, diabetes mellitus, alopecia totalis
Kufor Rakeb 606693	Rare	AR	ATP13A2/PARK9 (1p36.13)	Cation-transporting ATPase 13A2	Lysosomal dysfunction	Parkinsonism, supranuclear gaze palsy, dystonia	Cognitive decline, visual hallucinations
Niemann–Pick type C 257220	<1/120,000 births	AR	NPC1/NPC2 (18q11.2)	Niemann–Pick C1 protein	Free cholesterol and glycosphingolipid accumulation	Ataxia	Hepatosplenomegaly, seizures
DRPLA 125370	0.2–0.7/100,00 in Japanese population	AD	ATN-1 (12p13.31)	Atrophin-1	ATN-1 accumulation	Cerebellar ataxia, choreoathetosis	Dementia, seizures

AD, Autosomal dominant; AR, autosomal recessive; BPAN, beta-propeller protein-associated neurodegeneration; CoPAN, COASY protein-associated neurodegeneration; DRPLA, dentarubral–pallidylusian atrophy; FAHN, fatty acid hydroxylase-associated neurodegeneration; MPAN, membrane protein-associated neurodegeneration; PKAN, pantothenate kinase-associated neurodegeneration; PLAN, phospholipase A2-associated neurodegeneration.

**TABLE 27.6** Blood plasma tests associated with specific forms of genetically determined ataxia.

Disorder	Serum test
Ataxia telangiectasia	Alpha-fetoprotein (raised)
Ataxia with oculomotor apraxia type 2	Alpha-fetoprotein (raised)
Ataxia with vitamin E deficiency	Vitamin E (reduced)
Cerebrotendinous xanthochromatosis	Cholestinol (raised)
Oculomotor apraxia type 1	Albumin (reduced)

reported to be similarly affected; his mother, a maternal uncle and his son. On examination, there was evidence of impaired horizontal saccades and bilateral sixth cranial nerve palsies. He was dysarthric, although the majority of his speech was intelligible. Examination of the limbs demonstrated some generalized wasting and increased tone throughout. There was evidence of ataxia with both finger–nose and heel–shin testing. His gait was moderately broad based and ataxic, and he was unable to heel–toe walk.

Spinocerebellar ataxia 2 (SCA2) is caused by autosomal-dominant inheritance of CAG repeat expansions to the *ATXN2* gene. Typically, approximately 22 repeats are observed, with pathological signs and symptoms evident when >32 repeats are present. The number of repeats appear to be linked with age at onset, with those with >45 repeats frequently developing symptoms in their teens. *ATXN2* encodes the ataxin-2 protein whose function remains largely unknown; however, it appears to be predominantly present in cellular cytoplasm and may play a role in RNA processing through interaction with the ER. Treatment is supportive with encouragement of regular exercise to maintain muscle mass and a healthy weight. With progressive ataxia walking aids are frequently required, and computerized devices are often used to aid communication with more severe dysarthria.

#### 27.4.6 Spinocerebellar ataxia type 3: Machado–Joseph disease

##### Case report

A 50-year old woman was receiving ongoing follow-up in the neurology department. She had initially presented 10 years earlier with a subacute onset of progressive clumsiness and unsteadiness. This had progressed to involve speech and swallowing difficulties, such that she now had a modified, thickened diet. During this time, she had also reported blurring of her vision, which had later progressed to a horizontal diplopia. There were reports of disturbed sleep, with vocalizations and acting out of vivid dreams. Several other family members were affected, including her paternal grandfather, father, and several uncles. However, onset of their symptoms had been at least a decade later than hers. On examination, there was evidence of horizontal nystagmus and reported diplopia on the extremes of gaze. Examination of the limbs demonstrated increased tone throughout, most marked in the lower limbs, and a generalized hyperreflexia. Finger–nose testing in the upper limbs and heel–shin testing in the lower limbs found evidence of an intention tremor and dysmetria. Her gait was broad based, and she was unable to heel–toe walk.

SCA3 (previously known as Machado–Joseph disease) is caused by autosomal dominantly inherited mutations in the ataxin-3 gene (*ATXN3*). This is a trinucleotide repeat disorder caused by an expansion of the CAG repeat (normal range: 12–43 repeats). SCA3 also demonstrates anticipation, with those <75 repeats tending to develop symptoms in mid-adulthood where as individuals with ~80 repeats tend to present in their teenage years. The encoded ataxin-3 protein functions in removing ubiquitin from proteins undergoing degradation as part of the ubiquitin-proteasome system, freeing up the ubiquitin for future use. Pathological CAG expansions result in an abnormally long ataxin-3 enzyme, causing a loss of function with the combined protein, ubiquitin and ataxin-3 complex aggregating in the nucleus. The mechanism by which this results in neurodegeneration is not fully understood, but atrophy is initially prominent in the brainstem and cerebellum, with later involvement of the spinal cord.

#### 27.4.7 Spinocerebellar ataxia type 7

##### Case report

A 50-year old man described a 10-year history of progressive symptoms. These began with visual loss, initially involving his central vision, but progressing such that he was now registered as being blind. A few years later, he noticed difficulties

with coordination, a general clumsiness and increased number of falls. Over the preceding 12 months, he found that people were having greater difficulty understanding what he was saying, and he was having trouble swallowing solid food. On examination, there was evidence of some mild cognitive difficulties. Eye movement examination revealed generally slow saccades in both horizontal and vertical planes. There was loss of coordination in both upper and lower limbs, resulting in past-pointing, intention tremor, and dysidiadochokinesis. Five other family members had developed similar symptoms including the patient's older sister, his mother, and maternal grandfather.

SCA7 is caused by a trinucleotide repeat (CAG) expansion in the ATXN7 gene, resulting in production of a pathological form of the ataxin-7 protein. As with SCA6 this is thought to lead to nuclear accumulation of protein aggregates resulting in neuronal loss. Cerebral imaging demonstrates preferential atrophy of the cerebellum and pons, with more generalized atrophy as the disease progresses.

### 27.4.8 Ataxia-telangiectasia

#### Case report

A 10-year old boy was initially seen in the pediatric clinic at the age of 4 years with a 2-year history of balance difficulties. He had been born at term, at the end of a normal pregnancy. He had initially met all of his developmental milestones, but from the age of 2-years he had begun to fall more frequently and had difficulty with coordinated activities. Over this time period, he had had a number of hospital admissions due to respiratory and middle ear infections, many of which required prolonged courses of antibiotics. When he started school (aged 5 years), his teachers described a tendency for his gaze to flit between objects, often moving his head to be able to focus. Over the past 5 years the balance difficulties had progressed, such that he needed multiple walking aids to be able to mobilize short distances. On examination, there was evidence of enlarged blood vessels in the sclerae bilaterally, as well as on areas of sun exposed skin (telangiectasia). There was evidence of oculomotor apraxia (loss of co-ordinated head and eye movements) as well as upper and lower limb ataxia. He was unable to walk even short distances unaided. There was no reported family history of similar symptoms.

Ataxia-telangiectasia is caused by autosomal recessive inheritance of mutations in the ATM gene. The ATM protein is thought to be involved in regulating cellular responses to stress, including the repair of DNA in the context of double-strand breaks. Under normal circumstances the ATM protein would halt the cell cycle and recruit repair proteins to the site, facilitating a regulated repair process. In the presence of ATM mutations, these repair mechanisms are impaired, resulting in genomic instability and an increased risk of malignancy, especially leukemia and lymphoma. As well as genetic testing, other clinical investigations include cerebral imaging (cerebellar atrophy), serum alpha-fetoprotein (elevated) and serum immunoglobulins (reduced).

#### 27.4.8.1 Therapy of ataxia telangiectasia (AT)

Few disease-modifying therapies are available in the treatment of the hereditary ataxias, although treatments do exist for reversible forms, for example, replacement of vitamin E in vitamin E-deficient ataxia, replacement of fat-soluble vitamins and a low-fat diet in abetalipoproteinemia, and oral chenodeoxycholic acid to prevent accumulation of metabolites in cerebrotendinous xanthomatosis [31].

### 27.4.9 Potential future targets for molecular therapy

As several of the genetically determined forms of ataxia are caused by over or under gene expression rather than altered protein product, potential future therapies include the direct targeting of gene expression, such as in Friedreich's ataxia where there is reduced production of the frataxin protein. Another potential mechanism of treatment is the targeting of abnormal protein aggregates, for example, heat shock protein 70, involved in protein quality control pathways, is reduced in SCA7. Activators of this protein could potentially be utilized therapeutically to improve protein regulation and reduce aggregation. Oxidative stress has also been implicated in ataxia pathophysiology, where mitochondrial iron accumulation and impaired respiratory chain electron transport lead to free radical formation, oxidative stress, and DNA damage. Preclinical testing of antioxidant treatment has shown some promising results, although these are yet to progress to clinical trials. Therapies aimed at modulating calcium signaling pathways have also been proposed, for example, the abnormal ataxin-2 protein in SCA2 interacts with the inositol 1,4,5-triphosphate receptor resulting in increased intracellular release of calcium; enzymes that reduce level of the receptor substrate (inositol 1,4,5-triphosphate) have been shown to improve motor function in mice [29,32].



## 27.5 Ataxia: key learning points

- Several of the inherited ataxias are trinucleotide repeat disorders demonstrating anticipation of clinical symptom onset with successive generations.
- Clinical characteristics often involve impaired bulbar function (speech and swallow) as well as loss of limb coordination.
- Key processes likely to be involved in pathogenesis are DNA repair mechanisms, mitochondrial dysfunction and free radical accumulation, and loss of calcium homeostasis.
- No disease-modifying or curative therapies exist to date, although future treatment may involve gene therapy aimed at altering levels of gene expression.

## 27.6 Other movement disorders

### 27.6.1 Huntington's disease

#### Case report

A 45-year old gentleman presented with a 5-year history of depressive symptoms requiring ongoing treatment from the local psychiatry department. Over the past 2 years, family members had reported memory difficulties, particularly short-term memory impairment and increasing difficulty managing day-to-day living. This gentleman's GP had referred him to the neurology department after noticing persistent, fidgety limb and facial movements during a routine consultation. On examination, this gentleman was low in mood and had poor eye contact throughout the consultation. There was evidence of generalized choreiform movements, particularly involving the eyes, mouth, and limbs. He had difficulty performing tasks such as water pouring and writing a sentence. Further discussion with the patient and his accompanying family members revealed that five other family members were affected with similar symptoms. The patient's grandfather developed symptoms in his 70s, two of his four uncles developed a movement disorder in their mid-50s, and his cousin had presented to his local neurology department a few years earlier.

HD is an autosomal dominantly inherited neurodegenerative disorder and is most common among Caucasian individuals, with a prevalence of 5–7/100,000 [33]. Clinically it is characterized by motor abnormalities, such as chorea and dystonia, and neuropsychiatric symptoms, including cognitive decline and personality changes. It is caused by a CAG repeat expansion in the *Huntingtin* gene, encoding a prolonged polyglutamine repeat in the huntingtin protein, and resulting in a toxic gain of function. Repeat lengths of >36 are considered pathogenic with the mutations demonstrating anticipation, that is, accumulation of increased CAG repeats in successive generations results in an earlier age at onset, with CAG repeats of >55 typically associated with the juvenile form of the disease. Although the number of CAG repeats account for the majority of the variation in age at onset, a recent genome wide association study (GWAS) demonstrated the likely role of DNA repair pathways, and in particular mutations in the *FANL* gene [34]. It is the GABAergic striatal medium spiny neurons that are most vulnerable in HD, resulting in prominent atrophy of the caudate and putamen nuclei. The nuclear accumulation of mutant huntingtin, also known as intraneuronal nuclear inclusions, has been found in the brains of patients with HD, with protein misfolding and inadequate protein clearance suggested as potential pathogenic mechanisms.

#### 27.6.1.1 Treatment of Huntington's disease

The treatment of HD is symptomatic, with no disease-modifying or curative therapies available at present. The chorea observed in HD can be treated with dopamine-depleting agents such as tetrabenazine to varying effect. In those patients with dystonia, botulinum toxin injections may aid in reducing symptoms of abnormal posture and pain. Neuropsychiatric manifestations are common throughout the course of disease and atypical neuroleptics may be used to treat psychotic symptoms, while selective serotonin reuptake inhibitors (e.g., citalopram) can help in the management of depression, anxiety and obsessive–compulsive behavior.

### 27.6.2 Wilson's disease

Wilson's disease is caused by autosomal recessive inheritance of mutations in the *ATP7B* gene [35]. The *ATP7B* protein is responsible for incorporating copper into ceruloplasmin, and is among a group of proteins that utilize ATP to transfer metals into and out of cells. Mutated *ATP7B* results in changes to the copper-binding

domains of the ATP7B protein expressed in hepatocytes, resulting in hepatic copper accumulation with subsequent systemic release, leading to its deposition in the brain, kidneys, and cornea. In addition to genetic testing, the investigation of Wilson's disease typically includes the measurement of ceruloplasmin, and serum and urinary copper levels. Neuroimaging usually demonstrates changes to the basal ganglia with increased density visualized on CT head and hyperintensity on T2-weighted MRI. The clinical presentation is typically either hepatic in adolescence, or in later life with neuropsychiatric symptoms. Those with early-onset Wilson's disease tend to present with acute liver failure and/or hemolytic anemia. The presentation of neurological signs is predominantly extrapyramidal, including bradykinesia, tremor, and rigidity, while neuropsychiatric presentations include psychosis and depression. In addition, a classical feature of Wilson's is the appearance of Keyser–Fleischer rings, with copper deposition in the Descemet's membrane between the cornea and the sclera. Treatment is mainly by chelating agents to reduce copper levels, such as d-penicillamine, with hepatic transplantation considered in those with end-stage liver disease.

### 27.6.3 Neurodegeneration with brain iron accumulation

Neurodegeneration with brain iron accumulation is a rare (prevalence of  $<1/1,000,000$ ) neurodegenerative disorder characterized by extrapyramidal signs, intellectual disability, and iron deposition within the basal ganglia. Ten genetically determined forms have been identified: eight autosomal recessive, one autosomal dominant, and one X-linked dominant [36].

#### 27.6.3.1 Pantothenate kinase-associated neurodegeneration

##### *Case report*

*A 6-year old girl developed speech difficulties, followed by problems with balance and mobility resulting in a number of falls. Her parents also reported that her neck and limbs had begun to adopt unusual postures, and that she felt very “stiff” and “rigid” when they were helping her to dress in the morning. There were also reports of increasing difficulty with vision at night, and an apparent reduction in peripheral vision, with a tendency to bump into furniture at home. On examination, there was marked dysarthria and evidence of significant weight loss. There was a generalized spasticity, particularly involving the lower limbs, with brisk reflexes and bilateral upgoing plantar responses. Cervical dystonia was evident, with abnormal posturing of the neck and shoulder-girdle region. A recent review by an ophthalmologist had confirmed evidence of retinal degeneration. There was no history of a similar disorder affecting immediate family members.*

Pantothenate kinase-associated neurodegeneration (PKAN) is an autosomal recessive disorder caused by a mutation in the *PANK2* gene, which encodes the pantothenate kinase 2 enzyme, responsible for the phosphorylation of components involved in the first key regulatory step of coenzyme A biosynthesis. Two forms of PKAN have been described: the classical form usually affects those  $<6$  years old with a predominant lower limb dystonia, upper motor neuron signs, and frequent falls. Visual problems include pigmentary retinopathy and abnormal pursuit and saccadic eye movements. General decline follows with speech and swallowing difficulties presenting toward the latter stages of disease. The atypical form tends to progress more slowly, and presentation is more heterogeneous. Symptoms are generally age-dependent, with adolescents experiencing higher levels of dystonia, while adults tend to be affected by symptoms of parkinsonism (bradykinesia and rigidity). PKAN has a classical appearance on MRI with T2-weighted imaging demonstrating globus pallidus hypointensity with an anteromedial-placed region of hyperintensity, known as the “eye of the tiger” sign. Treatment is generally symptomatic, with therapy aimed at reducing symptoms of dystonia and spasticity with the use of anticholinergics, baclofen, botulinum toxin injections, benzodiazepines, and DBS.

#### 27.6.3.2 Phospholipase A2-associated neurodegeneration

Phospholipase A2-associated neurodegeneration (PLAN), an autosomal recessive disorder caused by a mutation in the calcium-independent phospholipase A gene, *PLA2G6*, presents in early childhood with developmental delay. There are three phenotypes. The first is an infantile neuroaxonal dystrophy that typically presents between 6 months and 3 years of age with developmental regression, hypotonia with subsequent progression to spastic tetraparesis. An atypical neuroaxonal dystrophy presents in later childhood with slower progression, dystonia, and spastic tetraparesis. The third phenotype consists of a *PLA2G6*-related dystonia-Parkinsonism, which normally presents in late adolescence/early adulthood [37]. Affected individuals develop truncal hypotonia, ocular involvement (optic atrophy and nystagmus), and cerebellar atrophy on cerebral imaging.

### 27.6.3.3 Mitochondrial membrane protein-associated neurodegeneration

Mitochondrial membrane protein-associated neurodegeneration (MPAN) is an adult-onset, autosomal recessively inherited disorder, although autosomal-dominant inheritance has been reported in one family. It is caused by mutations in the *C19orf12* gene, the function of which is poorly understood but is likely to be involved in fatty acid metabolism. Proposed mechanisms of pathogenesis include impaired mitochondrial autophagy and increased apoptosis [38,39]. Symptoms include dystonia (mainly of the hands and feet), optic atrophy, corticospinal tract signs such as spasticity and extensor plantars, behavioral difficulties, and cognitive impairment. MPAN is slowly progressive, with lower motor neuron signs manifesting as the disease progresses. Neuroimaging demonstrates iron accumulation within the pallidum and substantia nigra.

### 27.6.3.4 Beta-propeller protein-associated neurodegeneration

Beta-propeller protein-associated neurodegeneration is caused by X-linked dominant inheritance of mutations to the *WDR45* gene, resulting in abnormalities in the beta-propeller protein [40]. Females normally exhibit a heterogeneous *WDR45* germline pathogenic variant, whereas males either have a hemizygous *WDR45* pathogenic variant or a partial deletion of *WDR45*. The *WDR45* gene normally encodes a WD repeat domain phosphoinositide-interacting protein 4 (WIPI4), which is thought to play a role in autophagy, although its exact role remains unclear. Clinical features include developmental delay, seizures, dystonia, and parkinsonism. Gait impairment is prominent with toe walking and a broad-based ataxic gait. Symptomatic treatment with levodopa therapy has demonstrated some early benefit, but the effect is usually short-lived.

### 27.6.3.5 Fatty acid hydroxylase-associated neurodegeneration

Fatty acid hydroxylase-associated neurodegeneration (FAHN) is caused by autosomal recessive mutations to the *FA2H* gene, which plays an essential role in myelin production and cell cycle regulation [41]. *FA2H* encodes a fatty acid 2-hydroxylase that localizes to the ER and requires iron as a cofactor. It is postulated that a reduction in 2-OH fatty acids leads to the abnormal white matter changes typically seen in FAHN [41]. Presentation is usually in the first decade of life with disorders of gait and frequent falls, progressing to spasticity, dystonia, and ataxia.

### 27.6.3.6 Coenzyme A synthetase protein-associated neurodegeneration

Coenzyme A synthetase (COASY) protein-associated neurodegeneration is caused by an inborn error of metabolism in coenzyme A metabolism. Inheritance is autosomal recessive and involves a missense mutation in the *COASY* (coenzyme A synthetase), which encodes for a bifunctional enzyme that catalyzes the final two steps of coenzyme A synthesis, the biosynthesis of which is essential to human cells [42]. Affected individuals present with cognitive difficulties and gait abnormalities in the first decade of life. Dystonia and dysarthria are common, with the “eye of the tiger” pattern seen on neuroimaging.

### 27.6.3.7 Neuroferritinopathy

Neuroferritinopathy is a rare monogenic autosomal-dominant disorder characterized by a mutation in the gene encoding the L chain of ferritin (*FtL*), with mutations resulting in a change to the C-terminal portion of the protein. This leads to a reduction in the physical stability of the protein and a wider, permeable quaternary channel, leading to a greater propensity of ferritin to precipitate and form toxic aggregates [43]. Neuroferritinopathy typically presents in the fourth to sixth decade of life with dystonia (action-induced and orofacial), parkinsonism, and chorea, with later-onset cognitive decline [44].

### 27.6.3.8 Aceruloplasminemia

Aceruloplasminemia is an autosomal recessive disorder characterized by low or absent ceruloplasmin. The *CP* single copy gene encodes ceruloplasmin, a multicopper ferroxidase that facilitates iron export from cells and oxidizes  $\text{Fe}^{2+}$  to  $\text{Fe}^{3+}$ , enabling ferric iron to bind to transferrin [45]. The ceruloplasmin protein is expressed as a glycosylphosphatidylinositol (GPI)-linked form in astrocytes, with mutant forms resulting in ferrous iron being unable to become oxidized when it enters the CNS with subsequent astrocytic accumulation. Iron accumulation occurs in both the brain and viscera, with low serum levels of copper and iron and high serum ferritin levels [45]. Symptoms manifest during childhood with systemic features including diabetes, macrocytic anemia, cognitive decline, and facial dystonia, with the treatment of choice being iron chelation.

### 27.6.3.9 Kufor Rakeb

A variant of NBIA, Kufor Rakeb is an autosomal recessively inherited disorder caused by a mutation in the *ATP13A2* gene (*PARK9*) that encodes for a lysosomal type 5 P-type ATPase [46,47]. *ATP13A2* protein consists of 10 transmembrane domains, with both termini oriented toward the cytosol, and likely involved in maintaining a pool of healthy, functioning mitochondria [48]. This disorder is characterized by early-onset extrapyramidal signs with a supranuclear gaze palsy, hypometric saccades and dystonia. Neuroimaging demonstrates atrophy affecting cerebral and subcortical tissue [37,49].

### 27.6.4 Niemann–Pick type C

#### Case report

A 12-year old boy was referred by his general practitioner following reports of clumsiness, and difficulties with coordination at school and home. He was also reported to have hearing problems and had been diagnosed with sensorineural hearing loss a few years earlier. In the months prior to the appointment his parents had also noticed some slurring of his speech, and coughing when drinking water. Systemic examination revealed hepatosplenomegaly, later confirmed with ultrasound examination. Neurological examination revealed impaired vertical eye movements, with particular difficulty initiating saccades. There was evidence of a cerebellar ataxia with dysmetria and intention tremor on finger–nose testing, and a broad based ataxic gait when mobilizing independently. There was no reported family history of similar symptoms.

An autosomal recessive disorder, Niemann–Pick type C is a lipid storage disorder resulting in accumulation of cholesterol and other lipids in the liver, spleen, and brain, leading to progressive neurodegeneration. Diagnostic testing includes genetic testing and Filipin staining of skin fibroblasts. Niemann–Pick type C is characterized by two mutations: *NPC1* (on chromosome 18q11.2), which accounts for 95% of cases, and *NPC2* (on chromosome 14q24.3) responsible for the remaining 5% of cases. Common presenting features include hepatosplenomegaly, while neurological symptoms include ataxia and impaired horizontal saccades. The rate of neurodegeneration in NPC varies with age at onset, with those with earlier onset disease progressing at a faster rate than those with later-onset symptoms. Miglustat (reversible inhibitor of glycosphingolipid synthesis) has been demonstrated to slow disease progression, with greatest benefit in those with later-onset symptoms.

### 27.6.5 Dentarubral–pallidolusian atrophy

An autosomal dominantly inherited neurodegenerative disorder, dentarubral–pallidolusian atrophy (DRPLA) involves a combination of neurological and psychological signs and symptoms. These include a progressive dementia, seizures, and disorders of movement including myoclonus, chorea, and ataxia [50]. Those with early-onset DRPLA tend to develop progressive myoclonic epilepsy with dementia, whereas those with later-onset forms develop cerebellar ataxia, choreoathetosis, and dementia. DRPLA is caused by a trinucleotide (CAG) repeat expansion (>48 repeats being pathogenic) in the atrophin-1 (*ATN-1*) gene. *ATN-1* mutations involve a pathogenic gain of function, resulting in toxic neuronal accumulation of ATN-1. Inheritance of DRPLA mutations demonstrates both anticipation and somatic mosaicism, whereby the same individual has two genetically distinct population of cells [50].

## 27.7 Conclusion

The large number of disease-causing genes identified in movement disorders to date reflects the broad and complex nature of these disorders. While several of these genes have provided a mechanism for improved understanding of the pathology and biochemical disturbance observed in these conditions, further work is still required. In the majority of cases, treatment remains supportive and symptomatic, rather than disease-modifying or curative. Improved cellular and systemic understanding of disorder pathology should aid in drug-screening processes and identification of novel therapeutic targets.

## References

- [1] Farrer MJ. Genetics of Parkinson disease: paradigm shifts and future prospects. *Nat Rev Genet* 2006;7(4):306–18.
- [2] Kragh CL, Ubhi K, Wyss-Corey T, Masliah E. Autophagy in dementias. *Brain Pathol* 2012;22(1):99–109.
- [3] Lee S, Imai Y, Gehrke S, Liu S, Lu B. The synaptic function of LRRK2. *Biochem Soc Trans* 2012;40(5):1047–51.



- [4] Zimprich A, Benet-Pagès A, Struhal W, Graf E, Eck SH, Offman MN, et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *Am J Hum Genet* 2011;89(1):168–75.
- [5] Lücking CB, Dürr A, Bonifati V, Vaughan J, De Michele G, Gasser T, et al. Association between early-onset Parkinson's disease and mutations in the parkin gene. *New Engl J Med* 2000;342(21):1560–7.
- [6] Klein C, Westenberger A. Genetics of Parkinson's disease. *Cold Spring Harbor Perspect Med* 2012;2(1):a008888.
- [7] Pickrell AM, Youle RJ. The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. *Neuron* 2015;85(2):257–73.
- [8] Facheris M, Strain K, Lesnick T, Andrade M, Bower J, Ahlskog J, et al. UCHL1 is associated with Parkinson's disease: a case-unaffected sibling and case-unrelated control study. *Neurosci Lett* 2005;381(1–2):131–4. Available from: <<http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/722/CN-00813722/frame.html>>.
- [9] Bonifati V. Genetics of Parkinson's disease—state of the art, 2013. *Parkinsonism Relat Disord* 2014;20:S23–8.
- [10] Trinh J, Farrer M. Advances in the genetics of Parkinson disease. *Nat Rev Neurol* 2013;9(8):445–54.
- [11] Abou-Sleiman PM, Healy DG, Quinn N, Lees AJ, Wood NW. The role of pathogenic DJ-1 mutations in Parkinson's disease. *Ann Neurol* 2003;54(3):283–6.
- [12] Aharon-Peretz J, Badarny S, Rosenbaum H, Gershoni-Baruch R. Mutations in the glucocerebrosidase gene and Parkinson disease: phenotype–genotype correlation. *Neurology* 2005;65(9):1460–1.
- [13] Kalia LV, Lang AE. Parkinson's disease. *Lancet* 2015;386(9996):896–912. (London, England).
- [14] Sutton JP, Couldwell W, Lew MF, Mallory L, Grafton S, DeGiorgio C, et al. Ventroposterior medial pallidotomy in patients with advanced Parkinson's disease. *Neurosurgery* 1995;36(6):1112–17.
- [15] Simuni T, Jaggi J, Mulholland H, Hurtig H, Colcher A, Siderowf A, et al. Bilateral stimulation of the subthalamic nucleus in patients with Parkinson disease: a study of efficacy and safety. *J Neurosurg* 2002;96(4):666–72. Available from: <<http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/928/CN-00379928/frame.html>>.
- [16] Albanese A, Bhatia K, Bressman SB, Delong MR, Fahn S, Fung VS, et al. Phenomenology and classification of dystonia: a consensus update. *Mov Disord* 2013;28(7):863–73.
- [17] Peall KJ, Smith DJ, Kurian MA, Wardle M, Waite AJ, Hedderly T, et al. SGCE mutations cause psychiatric disorders: clinical and genetic characterization. *Brain* 2013;136(Pt 1):294–303.
- [18] Heron SE, Dibbens LM. Role of PRRT2 in common paroxysmal neurological disorders: a gene with remarkable pleiotropy. *J Med Genet* 2013;50(3):133–9.
- [19] Grabowski M, Zimprich A, Lorenz-Depiereux B, Kalscheuer V, Asmus F, Gasser T, et al. The epsilon-sarcoglycan gene (SGCE), mutated in myoclonus-dystonia syndrome, is maternally imprinted. *Eur J Hum Genet* 2003;11(2):138–44.
- [20] Niethammer M, Carbon M, Argelan M, Eidelberg D. Hereditary dystonia as a neurodevelopmental circuit disorder: evidence from neuroimaging. *Neurobiol Dis* 2011;42(2):202–9.
- [21] Guehl D, Cuny E, Ghorayeb I, Michelet T, Bioulac B, Burbaud P. Primate models of dystonia. *Prog Neurobiol* 2009;87(2):118–31.
- [22] Vidailhet M, Grabli D, Roze E. Pathophysiology of dystonia. *Curr Opin Neurol* 2009;22(4):406–13.
- [23] Albanese A, Asmus F, Bhatia KP, Elia AE, Elibol B, Filippini G, et al. EFNS guidelines on diagnosis and treatment of primary dystonias. *Eur J Neurol* 2011;18(1):5–18.
- [24] Vidailhet M, Vercueil L, Houeto JL, Krystkowiak P, Benabid AL, Cornu P, et al. Bilateral deep-brain stimulation of the globus pallidus in primary generalized dystonia. *N Engl J Med* 2005;352(5):459–67.
- [25] Jayadev S, Bird TD. Hereditary ataxias: overview. *Genet Med* 2013;15(9):673–83.
- [26] McCullough SD, Grant PA. Histone acetylation, acetyltransferases, and ataxia—alteration of histone acetylation and chromatin dynamics is implicated in the pathogenesis of polyglutamine-expansion disorders. *Adv Protein Chem Struct Biol* 2010;79:165–203.
- [27] Orr HT, Zoghbi HY. Trinucleotide repeat disorders. *Annu Rev Neurosci* 2007;30:575–621.
- [28] Li X, Liu H, Fischhaber PL, Tang TS. Toward therapeutic targets for SCA3: insight into the role of Machado-Joseph disease protein ataxin-3 in misfolded proteins clearance. *Prog Neurobiol* 2015;132:34–58.
- [29] Matilla-Duenas A, Ashizawa T, Brice A, Magri S, McFarland KN, Pandolfo M, et al. Consensus paper: pathological mechanisms underlying neurodegeneration in spinocerebellar ataxias. *Cerebellum* 2014;13(2):269–302.
- [30] Hekman KE, Gomez CM. The autosomal dominant spinocerebellar ataxias: emerging mechanistic themes suggest pervasive Purkinje cell vulnerability. *J Neurol Neurosurg Psychiatry* 2015;86(5):554–61.
- [31] Ramirez-Zamora A, Zeigler W, Desai N, Biller J. Treatable causes of cerebellar ataxia. *Mov Disord* 2015;30(5):614–23.
- [32] Bushart DD, Murphy GG, Shakkottai VG. Precision medicine in spinocerebellar ataxias: treatment based on common mechanisms of disease. *Ann Transl Med* 2016;4(2):25.
- [33] Walker FO. Huntington's disease. *Lancet* 2007;369(9557):218–28.
- [34] Bettencourt C, Hensman-Moss D, Flower M, Wiethoff S, Brice A, Goizet C, et al. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann Neurol* 2016;79(6):983–90.
- [35] Roberts EA, Schilsky ML. Diagnosis and treatment of Wilson disease: an update. *Hepatology* 2008;47(6):2089–111.
- [36] Hogarth P. Neurodegeneration with brain iron accumulation: diagnosis and management. *J Mov Disord* 2015;8(1):1.
- [37] Gregory A, Hayflick S. Neurodegeneration with brain iron accumulation disorders overview. *GeneReviews®* [Internet], 2014 - [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov).
- [38] Deutschländer A, Konno T, Ross OA. Mitochondrial membrane protein-associated neurodegeneration. *Parkinsonism Relat Disord* 2017;39:1–3.
- [39] Gregory A, Hartig M, Prokisch H, Kmiec T, Hogarth P, Hayflick SJ. Mitochondrial membrane protein-associated neurodegeneration. *GeneReviews®* [Internet], 2014 - [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov).
- [40] Hayflick SJ, Kruer MC, Gregory A, Haack TB, Kurian MA, Houlden HH, et al. Beta-propeller protein-associated neurodegeneration: a new X-linked dominant disorder with brain iron accumulation. *Brain* 2013;136(6):1708–17.
- [41] Kruer MC, Gregory A, Hayflick SJ. Fatty acid hydroxylase-associated neurodegeneration. *GeneReviews* [Internet] Seattle (WA): University of Washington; 1993–2011.



- [42] Dusi S, Valletta L, Haack TB, Tsuchiya Y, Venco P, Pasqualato S, et al. Exome sequence reveals mutations in CoA synthase as a cause of neurodegeneration with brain iron accumulation. *Am J Hum Genet* 2014;94(1):11–22.
- [43] Levi S, Cozzi A, Arosio P. Neuroferritinopathy: a neurodegenerative disorder associated with L-ferritin mutation. *Best Pract Res Clin Haematol* 2005;18(2):265–76.
- [44] Crompton DE, Chinnery PF, Fey C, Curtis ARJ, Morris CM, Kierstan J, et al. Neuroferritinopathy: a window on the role of iron in neurodegeneration. *Blood Cells Mol Dis* 2002;29(3):522–31.
- [45] Levi S, Finazzi D. Neurodegeneration with brain iron accumulation: update on pathogenic mechanisms. *Front Pharmacol* 2014;5.
- [46] Tonekaboni SH, Mollamohammadi M. Neurodegeneration with brain iron accumulation: an overview. *Iran J Child Neurol* 2014;8(4):1.
- [47] Hampshire DJ, Roberts E, Crow Y, Bond J, Mubaidin A, Wriekat A-L, et al. Kufor-Rakeb syndrome, pallido-pyramidal degeneration with supranuclear upgaze paresis and dementia, maps to 1p36. *J Med Genet* 2001;38(10):680–2.
- [48] Grünwald A, Arns B, Seibler P, Rakovic A, Münchau A, Ramirez A, et al. ATP13A2 mutations impair mitochondrial function in fibroblasts from patients with Kufor-Rakeb syndrome. *Neurobiol Aging* 2012;33(8):1843-e1.
- [49] Schneider SA, Paisan-Ruiz C, Quinn NP, Lees AJ, Houlden H, Hardy J, et al. ATP13A2 mutations (PARK9) cause neurodegeneration with brain iron accumulation. *Mov Disord* 2010;25(8):979–84.
- [50] Wardle M, Morris HR, Robertson NP. Clinical and genetic characteristics of non-Asian dentatorubral-pallidoluysian atrophy: a systematic review. *Mov Disord* 2009;24(11):1636–40.



# Molecular pathology in neuropsychiatric disorders

*George Kirov and Elliott Rees*

Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, United Kingdom

## 28.1 Introduction

The major psychiatric disorders schizophrenia (SZ) and manic depressive disorder [also known as bipolar affective disorder type I (BPI)] affect up to 1% of the general population. They have long been recognized to have a strong genetic component. This has been confirmed by family, twin, and adoption studies (reviewed by, e.g., [1,2]). It has just been over 100 years since the first systematic study by Ernst Rüdin showed that the risk for SZ is increased in the relatives of patients, and subsequent studies confirmed that this risk is at least 10% in their siblings and children (e.g., [3,4]). Twin studies consistently showed a higher concordance in monozygotic (MZ) versus dizygotic (DZ) twins in both SZ and BPI [2,5], but approximately 50% concordance in MZ twins also indicated incomplete penetrance and the role of environmental factors. As upbringing and family environment was believed to cause the main psychiatric disorders, large adoption studies were conducted, but they also demonstrated that what contributed to the risk was primarily genetic risk among the biological parents [1,4]. Twin and family studies allow researchers to estimate the heritability of these disorders. This has been estimated at around 80% from twin studies of SZ [6] or 60% from population-based family studies [4]. Similar rates were also found for BPI [4,7]. The genetic susceptibility to develop these disorders is conferred by genetic factors that are either specific to one disorder or shared by both of them [2].

Despite the clear evidence that genetic factors are predominantly responsible for the development of SZ and BPI, there were almost no specific genetic findings until only a few years ago. Early linkage studies were highly inconsistent, and association studies suffered from a nearly universal lack of replication. It was becoming clear that these disorders were highly polygenic, had a complex inheritance, and that any genetic factor had incomplete penetrance. The picture changed from around 2008 after the introduction of SNP microarrays and later by high-throughput sequencing (HTS) technologies.

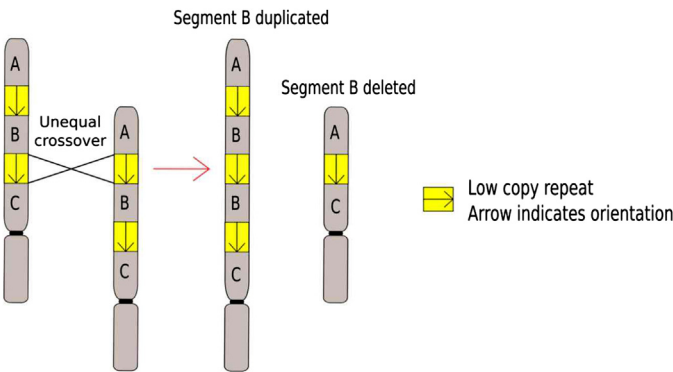
## 28.2 Copy number variation in psychiatric disorders

Copy number variations (CNVs) are structural alterations to chromosomes of >1000 bases in length [8]. The most extensively studied classes of CNV in psychiatric disorders are large deletions and duplications of >100 kb. Most of these CNVs are formed through nonallelic homologous recombination [9], facilitated by the presence in the human genome of repetitive DNA elements, low copy repeats, that cause chromosomes to misalign during recombination, leading to unequal chromosomal crossover that can result in deleted or duplicated DNA segments [10] (Fig. 28.1).

CNVs created via this mechanism can mutate relatively frequently and have similar breakpoints. They are called recurrent CNV [11]. Due to their relatively higher frequency, it has been possible to accumulate sufficient numbers of CNV observations to obtain strong statistical significance for their association with disease. Other

types of CNV can result from errors in DNA break repair and DNA replication, leading to unique genomic locations, therefore termed nonrecurrent CNV [9]. The first CNV shown to increase risk in developing SZ was the 22q11.2 deletion, also known as DiGeorge syndrome, or velocardiofacial syndrome [12], and for many years remained the only known genetic factor that increased risk of a major psychiatric disorder.

The discovery of the role of CNV in psychiatric disorders started in 2008 when two large independent studies identified the first CNV loci that increase risk for SZ [13,14]. Although recurrent, these CNVs were still relatively rare; therefore large samples of cases and controls were required to identify such associations. High-throughput genotyping with SNP microarrays and large-scale collaborations enabled the accumulation of tens of thousands of cases and controls. As of today, 13 CNVs are robustly associated with SZ (Table 28.1) with strong statistical support and replication [9,15–17].

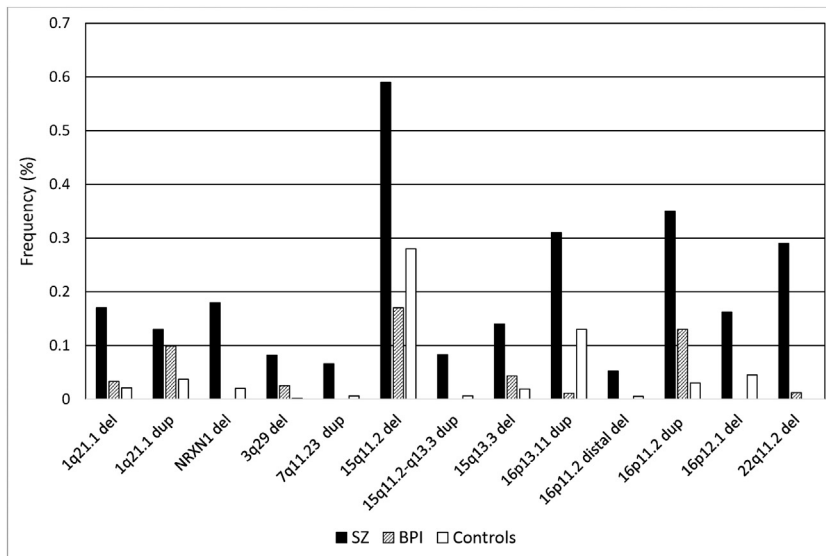


**FIGURE 28.1** Creation of CNVs via nonallelic homologous recombination (NAHR). Light-coloured segments indicate Low copy repeats (LCRs).

**TABLE 28.1** Rates of the most significant copy number variations (CNVs) in schizophrenia patients and controls.

Locus	Critical region (hg19)	Cases (N = 12,000–21,000) (%)	Controls (N = 20,000–81,000) (%)	OR (95% CI)	N genes in the interval	P-Value
1q21.1 del	chr1:146,6–147,4	0.17	0.021	8.35 (4.65–14.99)	11	$4.1 \times 10^{-13}$
1q21.1 dup	chr1:146,6–147,4	0.13	0.037	3.45 (1.92–6.20)	11	$9.9 \times 10^{-5}$
NRXN1 del	chr2:50,1–51,3	0.18	0.020	9.01 (4.44–18.29)	1	$1.3 \times 10^{-11}$
3q29 del	chr3:195,7–197,3	0.082	0.0014	57.65 (7.58–438.44)	21	$1.5 \times 10^{-9}$
7q11.23 dup	chr7:72,7–74,1	0.066	0.0058	11.35 (2.58–49.93)	28	$6.9 \times 10^{-5}$
15q11.2 del	chr15:22,8–23,1	0.59	0.28	2.15 (1.71–2.68)	4	$2.5 \times 10^{-10}$
15q11.2–q13.3 dup	chr15:24,8–28,4	0.083	0.0063	13.20 (3.72–46.77)	13	$5.6 \times 10^{-6}$
15q13.3 del	chr15:31,1–32,5	0.14	0.019	7.52 (3.98–14.19)	7	$4.0 \times 10^{-10}$
16p13.11 dup	chr16:15,5–16,3	0.31	0.13	2.30 (1.57–3.36)	8	$5.7 \times 10^{-5}$
16p11.2 del (distal)	chr16:28,8–29,1	0.052	0.005	20.6 (2.6–162.2)	11	$5.5 \times 10^{-5}$
16p11.2 dup	chr16:29,6–30,2	0.35	0.030	11.52 (6.86–19.34)	33	$2.9 \times 10^{-24}$
16p12.1 del	chr16:21.9–22.4	0.162	0.045	3.3 (1.61–7.05)	8	$3.4 \times 10^{-4}$
22q11.2 del	chr22:19,0–20,3	0.29	0.00	NA (28.27– $\infty$ )	40	$4.4 \times 10^{-40}$

The data are according to [15–17]. CNV coordinates are according to build hg19.



**FIGURE 28.2** Frequencies of selected CNVs in patients with SZ, BPI, and controls. BPI, Bipolar affective disorder type I; SZ, schizophrenia. Source: Data on BPI are from Green et al. (2016) [19].

This list is most likely incomplete and will be expanded as more samples and higher resolution studies become available [17,18].

Given the overlap of genetic risk between SZ and BPI [4], an obvious line of investigation was to check the involvement of these CNVs in BPI. It is now clear that these CNVs play a smaller role in BPI [19], with some of the CNVs from the list in Table 28.1 being close to control frequencies among bipolar patients. Only the 16p11.2 duplication, 1q21.1 duplication, and 3q29 deletions appear to be associated with BPI, but with lower frequencies than those observed in SZ (Fig. 28.2). In fact, it is clear from the figure that the frequencies of all these CNVs are lower in BPI than in SZ.

All SZ CNVs in Table 28.1 are also risk factors for intellectual disability (ID), developmental delay, autism spectrum disorders (ASDs), and congenital malformations [20,21]. In addition, a set of 51 CNVs implicated in ID are found at significantly higher rates in SZ patients than in controls [16]. Research on CNVs in other psychiatric disorders is insufficient to draw conclusions, as the number of patients recruited is still too small. The exception is ASDs, which is not discussed in this chapter.

### 28.3 Cognitive function among copy number variation carriers

It has been known for a long time that patients suffering with SZ have reduced cognitive performance [22]. As the 13 CNV loci associated with SZ are also risk factors for developmental delay, the question arises as to whether a proportion of the cognitive deficit in SZ patients can be explained by the presence of CNVs. Indeed, carriers of these CNVs have been shown to have reduced cognitive performance [23,24]. Using data from the UK Biobank, we showed that even among high-functioning individuals from the general population, who carry one of these CNVs, there are subtle reductions in cognitive performance [23]. If reduced cognitive performance is one of the risk factors for SZ then this can be one mechanism that can explain the role of CNVs in SZ. BPI is a disorder with much more subtle cognitive dysfunction, which again fits well with the observation of a much lower role of CNVs in this disorder.

### 28.4 Penetrance of copy number variations

Patients, carriers, their relatives, and psychiatrists will start asking questions about the penetrance of these CNVs, that is, what is the risk for carriers to develop a disorder. Nearly all CNVs from the list in Table 28.1 have incomplete penetrance. We have calculated the penetrance for these CNVs (Table 28.2) using the established frequencies among disease groups and healthy controls, and assuming rates of these disease groups in the general population of 1% for SZ and 4% for people referred for genetic testing [25]. It is of course important to be aware



**TABLE 28.2** Penetrance of SZ-associated CNVs.

Locus	Penetrance (%)		
	SZ	ID	Total
1q21.1 del	5.2	35	40
1q21.1 dup	2.9	18	21
NRXN1 del	6.4	26	32
3q29 del	18	53	71
WBS dup	6.0	44	50
15q11.2 del	2.0	11	13
PWS/AS maternal dup	12.3	50	62
15q13.3 del	4.7	35	40
16p13.11 dup	2.2	8.4	11
16p12.1 del	3.4	11	14
16p11.2 distal del	2.6	23	26
16p11.2 dup	8.0	26	34
DiGeorge/VCFS del	12	88	100

95% confidence intervals are very large and are not shown. Penetrance might be lower if we assume lower population frequencies of SZ and ID; on the other hand, they will be higher if subtle phenotypes, such as impaired cognitive performance, are taken into account in estimating the penetrance.

*Adapted and updated from Kirov G, Rees E, Walters TJ, Escott-Price V, Georgieva L, Richards AL, et al. The penetrance of copy number variations for schizophrenia and developmental delay. Biol Psychiatry 2014;75(5):378–85.*

that these CNVs not only increase risk for psychiatric disorders, but also for autism, developmental delay, and medical comorbidities. The penetrance for any disorder is much higher than that for SZ or other psychiatric disorders, and all of them have to be considered in the genetic counseling process. Carrying a CNV from this list, especially the ones with a lower penetrance, such as 15q11.2 deletion and 16p13.1 duplication, is compatible with normal functioning. Using results from the UK Biobank [23], we showed that most carriers of such CNVs had no neuropsychiatric diagnoses and they probably consider themselves healthy. They had only marginally lower school grades and just over 25% of carriers of neurodevelopmental CNV had been at college or university [23]. Good knowledge of the risks of every recurrent CNV will be important for professionals who provide genetic counseling, and more work should be conducted to provide this information.

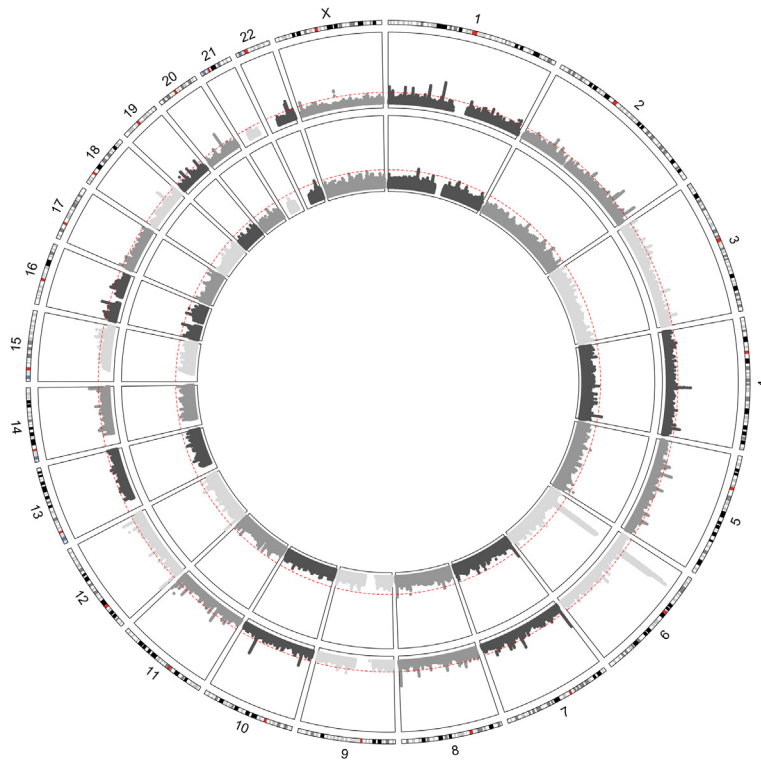
## 28.5 Results from genome-wide association studies

The same SNP microarrays that are used for CNV analysis provide information on hundreds of thousands of common genetic markers covering the whole genome. They are used for identifying association with common genetic variation and disease. Similar to the CNV analysis, tens of thousands of patients and controls needed to be genotyped in order to provide statistically robust associations that survive the massive multiple-testing problem.

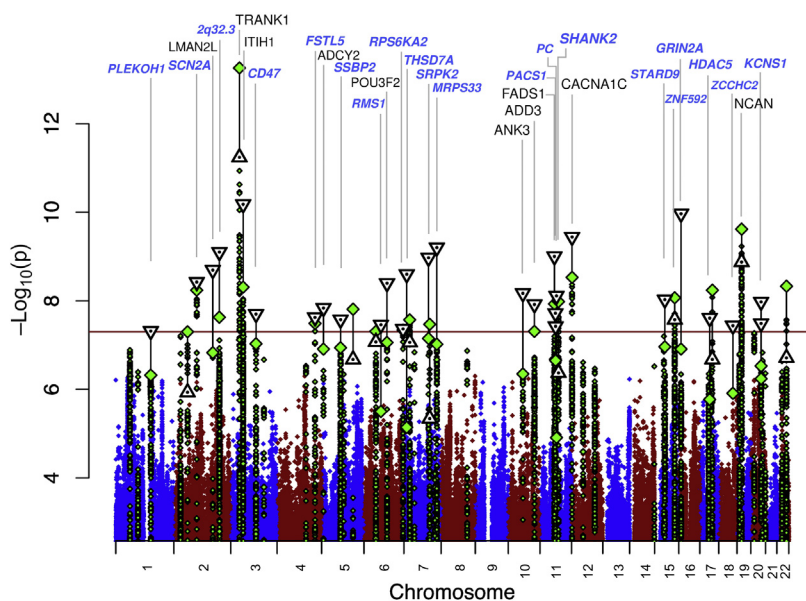
The latest *genome-wide association study* (GWAS) in SZ was conducted by Pardiñas et al., [26] on a new sample of 11,260 cases and 24,542 controls. Through meta-analysis with existing data (a total of 40,675 cases and 64,643 controls), the team identified 145 associated loci that exceeded the genome-wide correction threshold ( $P < 5 \times 10^{-8}$ ) for association with SZ. Loss-of-function (LoF)-intolerant genes were particularly enriched for association. The most strongly associated gene set constituted the targets of the fragile X mental retardation protein (FMRP). Other notable associations were with genes related to calcium signaling and the 5-HT<sub>2C</sub> receptor complex. These findings highlight molecules of known and potential therapeutic relevance to SZ. Another intriguing finding supports an old speculation about a link between the immune system and SZ, with the strongest genome-wide association at the extended major histocompatibility complex region on chromosome 6. The signal

has been narrowed down to alleles of the complement component 4 (C4) genes that generate widely varying levels of C4A and C4B expression in the brain [27]. These authors point out that excessive complement activity can contribute to the development of SZ by reducing the numbers of synapses in the brain (Fig. 28.3).

GWAS results on *bipolar disorder* are also available on very large samples [28]. A consortium performed the largest GWAS till date, including 20,352 cases (mostly BPI,  $n = 14,879$ ) and 31,358 controls of European descent, with follow-up analysis of 822 variants at loci with  $P < 1 \times 10^{-4}$  in an independent sample of 9412 cases and 137,760 controls. In the combined analysis, 30 loci achieved genome-wide significance (Fig. 28.4). These include neurotransmitter receptors (*GRIN2A*), ion channels and transporters (*CACNA1C*, *SCN2A*, *SLC4A1*), synaptic



**FIGURE 28.3** Circular Manhattan plot for GWAS in SZ. The inner circle shows the results from the new sample which includes patients recruited in the United Kingdom who take the antipsychotic Clozapine, while the outer circle represents the results from the full dataset. GWAS, Genome-wide association study; SZ, schizophrenia. Source: Figure produced by Antonio Pardiñas, first author of Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet 2018;50:381–9 paper, for the purpose of this chapter.



**FIGURE 28.4** Manhattan plot for the GWAS results on bipolar disorder. GWAS, Genome-wide association study. Source: Adapted from the data reported by Stahl E, Breen G, Forstner A, McQuillin A, Ripke S., Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, et al. Genome-wide association study identifies 30 loci associated with 1 bipolar disorder. BioRxiv 2018;173062; <https://doi.org/10.1101/173062>. 24/01/2018. In italics are shown the newly identified loci. [Special thanks to Eli Stahl (first author of that paper) for providing the image for the purpose of this chapter.]

components (*RIMS1*, *ANK3*), immune and energy metabolism components, and multiple potential therapeutic targets for mood stabilizer drugs.

There is substantial overlap between the genetics of SZ with the genetics of bipolar disorder, as shown in epidemiological studies that suggested substantial increases in the rates of SZ in first-degree relatives of bipolar probands, and vice versa [4]. Eight of the 30 associated loci in the GWAS (Fig. 28.4) also harbor SZ associations. The authors confirm a highly significant genetic correlation between bipolar disorder and SZ—the LD-score regression estimated genetic correlation was  $r_g = 0.70$ ,  $s.e. = 0.020$ .

A Psychiatric Genetics Consortium GWAS meta-analysis of 135,458 cases and 344,901 controls identified 44 risk loci for *major depression* [29]. Considering the larger numbers involved, it appears that this very common psychiatric disorder has relatively fewer clear susceptibility loci or that they have smaller effect sizes. The heritability of depression estimated from twin or family studies is also much lower, compared to SZ or bipolar disorder.

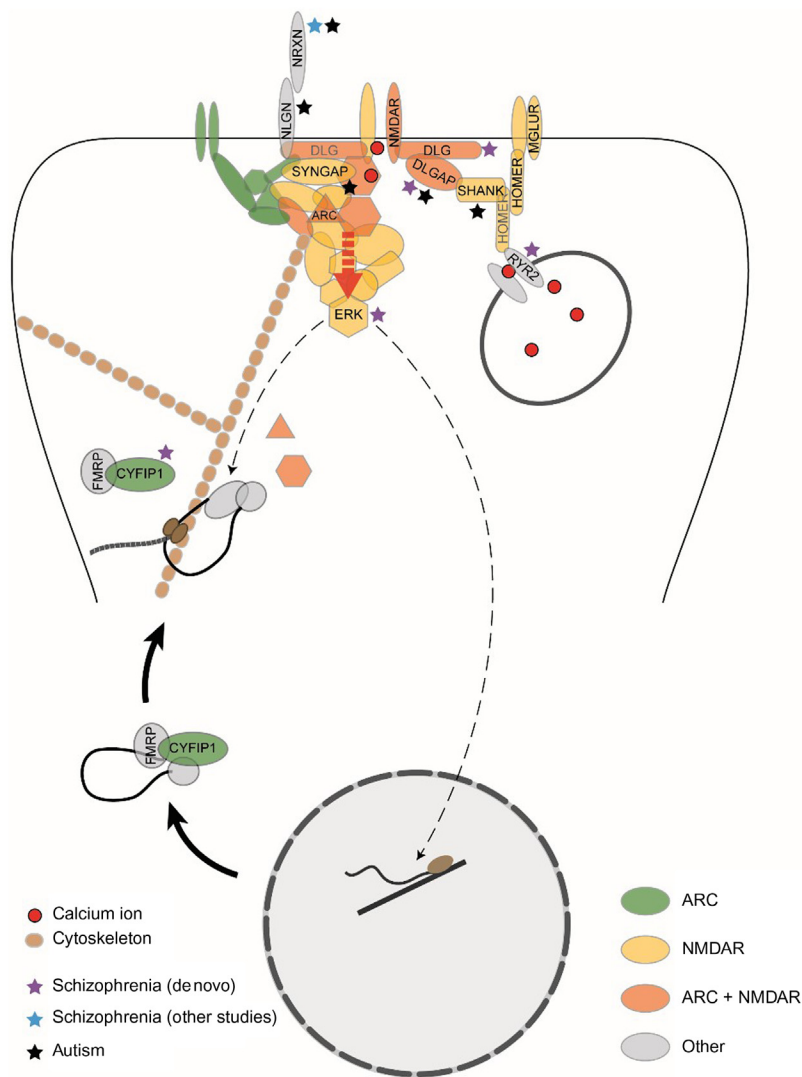
## 28.6 High-throughput sequencing studies

HTS studies are primarily designed to investigate ultrarare or de novo single-nucleotide variants (SNVs) and indels. These rare classes of mutation are almost always missed by SNP-genotyping arrays, and they are likely to be significant risk factors for disorders associated with reduced fecundity [30]. Most HTS studies of psychiatric disorders till date have sequenced protein coding regions of the genome, which is known as exome sequencing. In 2011 the first exome sequencing studies of SZ were published [31,32], which were small, exploratory, and inevitably underpowered. However, the rapid reduction in sequencing costs in recent years has led to larger and better powered studies. The largest sequencing study of de novo mutation in SZ did not find a significant excess of de novo mutations at the exome-wide level [33], which contrasts with findings from ID and ASD where an excess of LoF de novo mutations are shown [34–36]. However, de novo mutations in patients with SZ have been shown to be enriched among specific sets of genes. For example, the activity-regulated cytoskeleton-associated protein (ARC) and N-methyl-D-aspartate receptor (NMDAR) synaptic protein complexes were associated with de novo nonsynonymous and LoF mutations, respectively [33]. These findings replicated earlier associations made between these genes and CNVs [37]. Multiple studies have also shown SZ de novo mutations to be enriched in genes involved in chromatin remodeling, suggesting a role for epigenetic regulation in SZ pathogenesis [38,39]. Interestingly, genes affected by de novo mutations in patients with ASD and ID are also enriched for de novo mutations in SZ [33], supporting the evidence that these disorders share a genetic etiology.

Case–control sequencing studies of SZ have also shown rare coding alleles to significantly contribute to SZ risk. Although initial findings did not indicate an exome-wide excess of rare SNVs/indels [40], larger studies have shown ultrarare (defined as mutations observed in a single individual, and never in large population cohorts, e.g., ExAC <http://exac.broadinstitute.org/>) protein-altering mutations are more common in patients with SZ than controls [41]. This mutational excess observed in SZ was refined to 3388 genes that are expressed in neurons. Additional case–control sequencing studies of SZ have also shown a significant enrichment of rare LoF mutations in 3488 genes that are known to be intolerant of protein-truncating mutation, which is even stronger in SZ patients with comorbid ID [42].

Despite the strong evidence for a role of rare coding variants in SZ, only two genes have so far been associated with rare variants with robust statistical evidence. The enrichment of LoF alleles in *SETD1A*, which encodes a component of a histone methyltransferase complex, was discovered by combining case–control and trio sequencing data [43], indicating that joint analysis of inherited and de novo mutation can increase power to detect novel associations. The second single gene associated with SZ involves LoF mutations in *RBM12* (RNA-binding-motif protein 12), which was found to segregate with disease in an Icelandic family containing multiple members with psychosis [44].

The role of rare and de novo SNVs/indels in BPI is less clear, as they have been examined in much smaller samples than the studies of SZ. It therefore remains to be seen whether these rare classes of mutation have a significant impact on BPI risk. However, an initial sequencing study of BPI reported an enrichment of de novo mutations in genes that are intolerant of deleterious mutation [45]. In addition, a study which performed exome sequencing in multiplex bipolar families found nominally significant evidence that genes affected by rare mutations that segregated with disease were associated with BD in independent case–control sequencing data [46]. These results suggest that the role of rare/de novo SNVs/indels in BD warrants further exploration.



**FIGURE 28.5** Proteins in the postsynaptic density implicated by CNV and sequencing studies. Source: Figure produced and adapted by Pocklington AJ, Rees E, Walters JTR, Han J, Kavanagh DH, Chambert KD, et al. Novel findings from CNVs implicate inhibitory and excitatory signaling complexes in schizophrenia. *Neuron* 2015;86(5):1203–14.

## 28.7 Pathway analysis

Inferring disease biology from CNV and SNP associations is difficult as they often overlap multiple genes. A widely adopted approach is to test specific gene sets for enrichment. These sets of genes are usually grouped by common biological functions or expression profiles. CNVs in patients diagnosed with SZ have been shown to be enriched for genes encoding members of the postsynaptic density protein complex [37,47]. This association is largely driven by genes belonging to NMDAR and neuronal ARC complexes [37] (Fig. 28.5). Independent support for the involvement of these postsynaptic gene sets was obtained from case–control CNV studies [17,47,48] as well as exome sequencing studies of rare and de novo indel and point mutations [33,40]. These findings suggest disruption of glutamatergic signaling is involved in SZ pathogenesis. Additional gene sets associated with SZ CNVs include targets of the FMRP [17,26,48], genes involved in neuronal calcium channel signaling [26,48], and components of GABA<sub>a</sub> receptor complexes [47].

## 28.8 Conclusions

The last 10 years have witnessed the identification of a number of CNVs as strong risk factors for SZ and other neuropsychiatric disorders. It now appears that the same genetic factors can lead to different disorders, while many carriers appear asymptomatic. SZ patients carrying known risk CNVs also have an excess of common risk



alleles [49], suggesting that a combination of common and rare risk factors influences disease risk. It is likely that larger studies will uncover additional neuropsychiatric risk CNVs, and the PGC CNV group is currently working on analyses of disorders that have been less well-studied compared with SZ, such as BD, attention deficit hyperactivity disorder (ADHD), and posttraumatic stress disorder. The contribution from smaller and complex structural variations, best discovered through whole-genome sequencing, to neuropsychiatric risk remains largely unexplored. As more samples are sequenced, future studies could reveal these mutations to be additional neuropsychiatric risk factors.

## References

- [1] Kendall KM, Kirov G, Owen MJ. Chapter 12.4: Genetics of schizophrenia. In: Sadock BJ, Ruiz P, Sadock VA, editors. Kaplan and Sadock's comprehensive textbook of psychiatry. 2017, Lippincott, Williams and Wilkins.
- [2] Craddock N, Sklar P. Genetics of bipolar disorder. *Lancet* 2013;381(9878):1654–62. Available from: [https://doi.org/10.1016/S0140-6736\(13\)60855-7](https://doi.org/10.1016/S0140-6736(13)60855-7).
- [3] Gottesman II. Schizophrenia genesis. The origin of madness. New York: WH Freeman; 1991.
- [4] Lichtenstein P, Yip BH, Björk C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009;373(9659):234–9. Available from: [https://doi.org/10.1016/S0140-6736\(09\)60072-6](https://doi.org/10.1016/S0140-6736(09)60072-6).
- [5] Cardno AG, Gottesman II. Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. *Am J Med Genet* 2000;97:12–17.
- [6] Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait. Evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2007;60:1187–92.
- [7] McGuffin P, Rijdsdijk F, Andrew M, Sham P, Katz R, Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* 2003;60(5):497–502.
- [8] Lee C, Scherera SW. The clinical context of copy number variation in the human genome. *Exp Rev Mol Med* 2010;12:e8. Available from: <https://doi.org/10.1017/S1462399410001390>.
- [9] Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* 2012;148(6):1223–41.
- [10] Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 1998;14(10):417–22.
- [11] Lupski JR. Genomic rearrangements and sporadic disease. *Nat Genet* 2007;14:S43–7.
- [12] Murphy KC, Jones LA, Owen MJ. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 1999;56(10):940–5.
- [13] International Schizophrenia Consortium (ISC). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008;455(7210):237–41.
- [14] Stefansson H, Rujescu D, Cichon S, Pietiläinen OPH, Ingason A, Steinberg S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature* 2008;455(7210):232–6.
- [15] Rees E, Walters JTR, Georgieva L, Isles AR, Chambert KD, Richards AL, et al. Analysis of copy number variations at 15 schizophrenia-associated loci. *Br J Psychiatry* 2014;204(2):108–14.
- [16] Rees E, Kendall K, Pardiñas AF, Legge SE, Pocklington A, Escott-Price V, et al. Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA Psychiatry* 2016;73(9):963–9.
- [17] CNV and Schizophrenia Working Groups of the Psychiatric Genomics, C. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* 2017;49(1):27–35.
- [18] Rees E, Walters JTR, Chambert KD, Isles AR, Chambert KD, O'Dushlaine C, et al. CNV analysis in a large schizophrenia sample implicates deletions at 16p12.1 and *SLC1A1* and duplications at 1p36.33 and *CGNL1*. *Hum Mol Genet* 2014;23(6):1669–76.
- [19] Green EK, Rees E, Walters JTR, Smith KG, Forty L, Grozeva D, et al. Copy number variation in bipolar disorder. *Mol Psychiatry* 2016;21(1):89–93.
- [20] Coe BP, Witherspoon K, Rosenfeld JA, van Bon BWM, Vulto-van Silfhout AT, Bosco P, et al. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 2014;46(10):1063–71.
- [21] Dittwald P, Gambin T, Szafranski P, Li J, Amato S, Divon MY, et al. NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits. *Genome Res* 2013;23(9):1395–409.
- [22] Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology* 1998;12(3):426–45.
- [23] Kendall KM, Rees E, Escott-Price V, Eion M, Thomas R, Hewitt J, et al. Cognitive performance among carriers of pathogenic copy number variants: analysis of 152,000 UK Biobank subjects. *Biol Psychiatry* 2017;82(2):103–10.
- [24] Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdóttir B, Morgen K, Arnarsdóttir S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* 2014;505(7483):361–6.
- [25] Kirov G, Rees E, Walters TJ, Escott-Price V, Georgieva L, Richards AL, et al. The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry* 2014;75(5):378–85.
- [26] Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* 2018;50:381–9.
- [27] Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. *Nature* 2016;530(7589):177–83. Available from: <https://doi.org/10.1038/nature16549>.



- [28] Stahl E, Breen G, Forstner A, McQuillin A, Ripke S, Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, et al. Genome-wide association study identifies 30 loci associated with 1 bipolar disorder. *BioRxiv* 2018;173062. Available from: <https://doi.org/10.1101/173062> 24/01/2018.
- [29] Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 2018;50:668–81.
- [30] Bundy H, Stahl D, MacCabe JH. A systematic review and meta-analysis of the fertility of patients with schizophrenia and their unaffected relatives. *Acta Psychiatr Scand* 2010;123(2):98–106.
- [31] Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, et al. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet* 2011;43(9):864–8.
- [32] Girard SL, Gauthier J, Noreau A, Xiong L, Zhou S, Jouan L, et al. Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat Genet* 2011;43(9):860–3.
- [33] Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature* 2014;506:179–84.
- [34] Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 2012;485(7397):242–5.
- [35] O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 2012;485(7397):246–50.
- [36] Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 2012;485(7397):237–41.
- [37] Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry* 2012;17(2):142–53.
- [38] McCarthy S, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y, et al. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol Psychiatry* 2014;19(6):652–8.
- [39] Xu B, Ionita-Laza I, Roos JL, Boone B, Woodrick S, Sun Y, et al. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* 2012;44(12):1365–9.
- [40] Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* 2014;506:185–90.
- [41] Genovese G, Fromer M, Stahl EA, Ruderfer DM, Chambert K, Landen M, et al. Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat Neurosci* 2016;19(11):1433–41. Available from: <https://doi.org/10.1038/nn.4402>.
- [42] Singh T, Walters JTR, Johnstone M, Curtis D, Suvisaari J, Torniainen M, et al. The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. *Nat Genet* 2017;49(8):1167–73. Available from: <https://doi.org/10.1038/ng.3903>.
- [43] Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat Neurosci* 2016;19(4):571–7.
- [44] Steinberg S, Gudmundsdottir S, Sveinbjornsson G, Suvisaari J, Paunio T, Torniainen-Holm M, et al. Copy number variation in schizophrenia in Sweden. *Mol Psychiatry* 2014;19(7):762–73.
- [45] Kataoka M, Matoba N, Sawada T, Kazuno AA, Ishiwata M, Fujii K, et al. Exome sequencing for bipolar disorder points to roles of de novo loss-of-function and protein-altering mutations. *Mol Psychiatry* 2016;21:885.
- [46] Goes FS, Pirooznia M, Parla JS, et al. Exome sequencing of familial bipolar disorder. *JAMA Psychiatry* 2016;73(6):590–7.
- [47] Pocklington AJ, Rees E, Walters JTR, Han J, Kavanagh DH, Chambert KD, et al. Novel findings from CNVs implicate inhibitory and excitatory signaling complexes in schizophrenia. *Neuron* 2015;86(5):1203–14.
- [48] Szatkiewicz JP, O'Dushlaine C, Chen G, Chambert K, Moran JL, Neale BM, et al. Copy number variation in schizophrenia in Sweden. *Mol Psychiatry* 2014;19(7):762–73. Available from: <https://doi.org/10.1038/mp.2014.40>.
- [49] Tansey KE, Rees E, Linden DE, Ripke S, Chambert KD, Moran JL, et al. Common alleles contribute to schizophrenia in CNV carriers. *Mol Psychiatry* 2016;21(8):1153. Available from: <https://doi.org/10.1038/mp.2015.170>.



# Targeted molecular therapy: the cancer paradigm

*Dyfud Mark Davies*

Department of Oncology, Singleton Hospital, Swansea, Wales, United Kingdom; Division of Cancer & Genetics, Institute of Medical Genetics, Cardiff University, Cardiff, United Kingdom

## 29.1 Introduction

Significant strides have been made in our understanding of the genetics of cancer, beginning with the initial identification of oncogenes and tumor suppressor genes and culminating in the genomic profiling of tumors being routinely available in clinical practice. This progress has also led to some remarkable successes in molecularly targeted therapy, with the development of agents that target genetically driven tumor dependencies and vulnerabilities. This chapter will discuss the broad principles underlying targeted therapy and the difficulties that have been encountered, using illustrative examples from a range of cancer types.

## 29.2 Oncogene addiction

The concept of “oncogene addiction,” where the survival of cancer cells is highly dependent on the activity of a single-gene product, underpins many molecularly targeted therapies [1]. The treatment of chronic myeloid leukemia (CML) with tyrosine kinase inhibitors (TKIs) is a paradigm for this approach. Here, a single dysregulated protein is a fundamental driver of the disease; the protein has a “druggable” gain of function, and this protein is expressed in [2] leukemic cells but not in normal cells [3].

In CML the “druggable” target arises from a gene fusion. In 1960 a small derivative chromosome was noted to be consistently seen in the bone marrow cells of CML patients [4]. This chromosome was named the “Philadelphia” (Ph) chromosome, after the city in which it was first reported. The Ph chromosome was shown to arise from a reciprocal translocation between chromosomes 9 and 22: t(9;22)(q34;q11) [5]. This translocation results in the juxtaposition of the Abelson 1 (*ABL1*) gene, located on the long arm of chromosome 9, with a gene called *BCR* for breakpoint cluster region, on the long arm of chromosome 22 [2]. The normal *ABL1* protein is a tyrosine kinase involved in a wide range of cellular processes, including regulation of cell growth, survival, and migration, interacting with several intracellular signaling pathways including the RAS/RAF/MEK pathway, the JAK2/STAT pathway, and the PI3K/mTOR pathway. The protein arising from the chimeric *BCR-ABL1* gene had tyrosine kinase activity, derived from *ABL1* but deregulated as a consequence of the translocation. The first TKI to be used in the treatment of CML was imatinib [6]. This compound was initially identified from a chemical library screen of inhibitors of protein kinase A but was then subsequently shown to be an inhibitor of multiple kinases including *ABL1*. Early phase trials of imatinib in CML showed dramatic responses, and a subsequent phase III trial demonstrated the superiority of imatinib to interferon, the standard treatment for CML at that time [7]. The introduction of imatinib revolutionized the treatment of CML, leading to a dramatically better outcome for patients and providing the one of the first examples of a molecularly targeted approach to cancer treatment.

Oncogene addiction has also been successfully exploited in the treatment of melanoma and nonsmall cell lung cancer (NSCLC). Over 80% of melanomas harbor a mutation in v-raf murine sarcoma viral oncogene homolog b1(BRAF) [8], a serine/threonine protein kinase that plays a crucial role in the mitogen-activated protein kinase (MAPK) signaling pathway, involved in cell growth, proliferation, survival, and differentiation. The most frequent BRAF mutation is a glutamic acid–base substitution for valine at codon 600 (BRAF V600E). This mutation causes constitutive activation of the MAPK pathway, which, in turn, drives tumor progression. The first drug developed to target BRAF V600E specifically was vemurafenib, a small molecule reversible inhibitor with specific affinity for the adenosine tri-phosphate (ATP)-binding pocket of BRAF V600E [9]. Early phase I trials of vemurafenib in melanoma showed unprecedented clinical response rates (over 50%). BRIM-3, a phase III trial of vemurafenib in patients with BRAF V600 mutation-positive metastatic melanoma, reported an improved progression-free survival (PFS) and overall survival (OS) for vemurafenib compared with dacarbazine [10]. A second mutant BRAF inhibitor, dabrafenib has comparable efficacy, and both drugs have been approved for use in advanced malignant melanoma harboring a BRAF V600 mutation [11].

In many cases, NSCLC exhibits addiction to mutation within the epidermal growth factor receptor (EGFR) gene. These mutations are found in approximately 15% of cases in Western populations. In the early 2000s EGFR TKIs were initially trialed in an unselected NSCLC patient population based on the high proportion of NSCLCs known to express EGFR [12]. The results from these studies were disappointing, but the analysis of tumor tissue from responders demonstrated a link between certain EGFR kinase domain mutations and response [13]. In phase II studies, which enrolled patients with NSCLC harboring these mutations, treatment with EGFR TKIs resulted in an objective response rate of 65%–78% and PFS of 8.9–9.7 months [14]. The importance of molecular stratification was highlighted in the landmark IPASS trial that showed patients who had tumors with *EGFR* sensitizing mutations had significantly better PFS when treated with the TKI, gefitinib, compared with chemotherapy, whereas those with wild-type *EGFR* had significantly worse PFS with gefitinib [15]. The success of EGFR TKIs led to efforts to identify other actionable targets in NSCLC. Anaplastic lymphoma kinase (ALK) rearrangements occur in approximately 5% of NSCLCs [16]. In 95% of cases, ALK is fused to EML4 through a translocation, leading to activation of ALK. Crizotinib, a first-generation multitargeted TKI that inhibits ALK, improved PFS compared to chemotherapy in the first-line setting. More potent and selective ALK TKIs (ceritinib, alectinib, and brigatinib) have been subsequently developed and approved for treatment of ALK-rearranged NSCLC. ROS1 rearrangements occur in 1%–2% of NSCLC patients [17]. The most frequent fusion partner is CD74 (40%–45%), but a larger number of fusion partners have been identified in ROS1-rearranged NSCLC than ALK-rearranged NSCLC. ROS1 is a tyrosine kinase that promotes survival and proliferation through downstream signaling via SHP-1/SHP-2, JAK/STAT, PI3K/AKT/mTOR, and MAPK/ERK pathways. ROS1 translocations lead to fusions of an intact ROS1 tyrosine kinase domain with partner genes. Crizotinib, which potently targets ROS1, has been approved for treatment of ROS1-rearranged NSCLC.

Despite deep and sometimes durable responses of cancer to targeted drugs resistance invariably occurs, often within a year of starting treatment. Resistance to targeted agents can be classified as intrinsic, adaptive, or acquired [18,19]. Some tumors exhibit intrinsic resistance and do not respond to initial treatment. In other patients, there is an initial response, but adaptive resistance occurs when the tumor cells undergo changes in cell functioning, allowing survival during therapy. Acquired resistance can arise from selection for preexisting genetic alterations within a heterogeneous cancer cell population and the acquisition of new alterations during treatment. Resistance can develop by mechanisms that are either (1) “on target,” which is direct target reactivation mutations, or (2) “off target,” including activation of upstream effectors or downstream bypass signaling pathways and engagement of adaptive survival mechanisms.

Restoring the biologic function of targeted oncoproteins is a critical mechanism by which cancer cells can overcome targeted therapy. Acquired resistance to imatinib in BCR–ABL1 fusion-positive CML can be driven by secondary mutations in the drug-binding site, the ATP-binding pocket in the catalytic domain [3]. In NSCLC, EGFR mutations most commonly involve a deletion in exon 19 (19 Del) or substitution of leucine with arginine at codon 858 in exon 21 (L858R) [20]. EGFR exon 20 insertion mutations are the third most common type of EGFR mutation encountered in NSCLC. These mutations are present in 4%–9% of EGFR-mutant NSCLCs and confer intrinsic resistance to first-generation EGFR TKIs. The most common mechanism of acquired resistance is a secondary single base substitution in EGFR exon 20, resulting in a T790M mutation [21]. The third-generation EGFR inhibitor osimertinib was designed to target the T790M clone with maintained activity against the original exon 19del and L858R mutations. Osimertinib initially gained US Food and Drug Administration (FDA) approval for patients with metastatic EGFR T790M-mutant NSCLC after progressing on first- or second-generation EGFR TKI, but it subsequently gained approval in the first-line treatment of EGFR-mutant lung cancer. A phase III clinical

trial comparing osimertinib to first-generation EGFR TKIs reported a significant improvement in PFS with osimertinib (18.9 vs 10.2 months) [22]. This demonstrates the benefit of the anticipation and targeting of common resistance mechanisms.

The initial use of BRAF inhibitors in *BRAF V600E*-positive melanoma demonstrated that responses to the single agent were often dramatic but of short duration, typically lasting 6–8 months [23]. Multiple resistance mechanisms have been described, which often lead to the restoration of MAPK signaling, amplification of the mutated BRAF allele, and activating mutations in *NRAS* or *NF1* [24–27]. Moreover, a considerable number of patients treated with single-agent BRAF inhibitors developed secondary skin cancers, including squamous cell carcinoma and keratoacanthoma, due to paradoxical activation of the MAPK pathway in BRAF wild-type cells [28]. In light of these findings, clinical trials were conducted treating patients with combined BRAF and MEK inhibition. A phase III study in BRAF V600-mutant melanoma patients showed the superiority of the BRAF inhibitor, dabrafenib, also the MEK inhibitor, trametinib, compared with dabrafenib alone. Patients in the combination arm had a median PFS of 11 months and OS of 25.1 months as compared with PFS of 8.8 months and OS of 18.7 months in those who only received dabrafenib [29]. Furthermore, the number of skin cancers was much lower in the combination arm compared with the dabrafenib only arm.

Adaptive responses by cells are a means of developing resistance to a drug, which does not require additional genetic events [30]. For example, an adaptive upregulation of NF- $\kappa$ B pathway appears to be a mediator of resistance to EGFR TKI treatment in NSCLC [31]. Identification of such adaptive responses may allow the rational selection of drug combinations designed to inhibit such early prosurvival adaptive responses.

---

### 29.3 Synthetic lethality

---

Synthetic lethality refers to a situation where the loss of either one of two genes has little or no effect, but the combination of the loss of both genes is lethal [32]. Within a cancer cell, a genetic alteration, such as a defect in a tumor suppressor gene, can cause a second gene to become essential for cell survival. Targeting of the product of this second gene might be lethal to cancer cells but relatively nontoxic to normal cells.

The most clinically advanced use of synthetic lethality involves targeting tumors harboring BRCA1/2 loss-of-function with poly(ADP-ribose) polymerase (PARP) inhibitors [33]. PARP1 and PARP2 sense single-stranded DNA breaks and other types of DNA damage and upregulate the DNA-damage response (DDR) [34]. The tumor suppressor genes, BRCA1 and BRCA2, encode proteins involved in the repair of double-stranded DNA breaks by homologous recombination (HR) [35]. Heterozygous mutation of BRCA1 and BRCA2 in the germ line leads to an increased risk of breast, ovarian, prostate, and other cancers [36]. Tumors arising in individuals with a germ-line BRCA1 or BRCA2 mutation often have an acquired “second” hit, a somatic loss-of-function mutation, in the corresponding wild-type BRCA allele and therefore have defective HR. The synthetic lethality between PARP inhibition and loss of BRCA function is thought to be related to an increased number of double-strand DNA breaks or collapsed replication forks induced by PARP inhibition or PARP trapping on DNA [37].

Early phase trials demonstrated deep and durable responses in BRCA1- or BRCA2-deficient cancers to the PARP inhibitor olaparib. These findings have been confirmed in phase III trials, and PARP inhibitors are now in clinical use in BRCA1- or BRCA2-deficient breast or ovarian cancer [37]. Acquired resistance to PARP inhibitors has been observed by mechanisms, which restore HR functioning of such reversion mutation in BRCA1 or BRCA2 leading to or alterations in other components of the DDR [38].

A synthetic lethal relationship between PARP inhibitors and another component of the HR system might exist. HR deficiency results in a characteristic mutation signature, and identification of the presence of this signature might predict response to PARP inhibitors when a mutation in a specific DDR gene has not been identified [39].

---

### 29.4 Histology agnostic treatment

---

Traditionally, cancer drug development has followed a well-established pathway with drugs being developed to treat tumor types defined by their presumed anatomical origin, for example, lung cancer or colorectal cancer. However, the recent approval of programmed cell death-1 (PD-1) inhibitory antibodies, pembrolizumab and the tropomyosin receptor kinase (TRK) inhibitor, larotrectinib for use in a histological agnostic manner has established a new paradigm. These agents can be used in any tumor type, provided a biomarker predicting response is present.



Microsatellite instability (MSI) is the biomarker that predicts response to pembrolizumab [40]. Microsatellites are repetitive sequences found throughout the human genome. These sequences are prone to accumulation of mutations, mainly due to slippage of polymerases during DNA synthesis. The mismatch repair (MMR) system is responsible for excising mismatched nucleotide that can result from polymerase misincorporation errors, following recombination and from chemical or physical damage to nucleotides. Mutations in the genes that encode MMR components can lead to defective MMR (dMMR) and an accumulation of mutations, which can be detected as MSI. Tumors with significant MSI are termed MSI high (MSI-H) [41]. MSI-H has been observed in multiple tumor types, including colorectal, gastric, endometrial, and ovarian. Approximately 15% of colorectal cancers exhibit MSI-H owing to either epigenetic silencing of *MLH1* or a germline mutation in one of the MMR genes *MLH1*, *MSH2*, *MSH6*, or *PMS2* [42]. When anti-PD-1 antibodies were first tested in patients with colorectal cancer, a response was only seen in MSI-H tumors [43]. It was hypothesized that dMMR led to an increase in tumor mutational burden, which could be recognized by the patient's immune system following immune checkpoint blockade, leading to dramatic responses to treatment. This hypothesis was tested in some trials, which showed a very much higher response rate to PD-1 antibodies in patients with MSI-H tumors.

In 2017 the FDA approved pembrolizumab for the treatment of adult and pediatric patients with unresectable or metastatic solid tumors that have been identified as having MSI-H or dMMR. This was the first FDA approval for cancer treatment based on a biomarker rather than the anatomical site of the tumor. Later in 2017 the FDA granted approval for nivolumab, another PD-1 inhibitory antibody, as a treatment for patients with MSI-H or dMMR metastatic colorectal cancer after progression on standard chemotherapy.

The histology agnostic indication in MSI-H and dMMR tumors was approved after pembrolizumab had already obtained approvals for multiple, histologically defined indications, such melanoma and NSCLC and the safety profiles of the drug had been well established. The trials that lead to the approval of larotrectinib were explicitly done in a tissue agnostic fashion in patients with tumors with neurotrophic receptor tyrosine kinase (NTRK) fusions. The NTRK genes *NTRK1*, *NTRK2*, and *NTRK3* encode the proteins TRKA, TRKB, and TRKC, respectively [44]. Recurrent chromosomal fusion events involving the carboxy-terminal kinase domain of TRK have been identified around 1% of all solid tumors that occur in children and adults. These fusions lead to over-expression of the chimeric protein and constitutively active downstream signaling.

The efficacy of targeting TRK fusions in the broad range of cancer types has been tested using a basket trial approach. Here patients are selected by the molecular pathology of their tumors rather than by histological classification, based on the organ of origin. Patients whose tumors contain the qualifying genomic alterations can be entered, irrespective of cancer type. A basket trial of larotrectinib, a highly selective TRK inhibitor, included patients of any age and with any tumor type if a TRK fusion was present and reported a response rate of 75% [45]. In November 2018 the FDA granted accelerated approval to larotrectinib for adult and pediatric patients with solid tumors that have a NTRK gene fusion.

This is a significant change of perspective in oncology drug development moving from seeking activity in a histology defined tumor group to taking a histology agnostic approach from the outset. This is likely to be a recurrent theme in the development of the next generation of cancer drugs, but *the* context in which a biomarker occurs will often remain of importance. For example, *BRAF V600* mutations occur in a range of tumor types in addition to melanoma, but the sensitivity that these mutations confer to vemurafenib varies significantly between tumor type. *BRAF V600* mutations are found in NSCLC, colorectal cancer, papillary thyroid cancer, cholangiocarcinoma, Langerhans cell histiocytosis, and Erdheim–Chester disease. In some of these tumor types the incidence of mutations *BRAF V600* mutations is around 50%, but in more than half the incidence of mutations is less than 5%. In NSCLC, *BRAF V600* is found in approximately 3% of cases. A basket trial of vemurafenib in patients in the broad range of cancer types, which contained with *V600* mutations, showed a high response rate in patients with NSCLC, Erdheim–Chester disease, and Langerhans cell histiocytosis; but for some tumor types, such as colon cancer, the *V600* mutations were associated with a low response rate [46].

---

## 29.5 Limitations of molecularly targeted therapy in cancer

---

There has been undoubted success with molecular targeted therapy in cancer, but there are also ongoing constraints. A recent randomized trial (the SHIVA study) found equivalent outcomes between patients with a range of tumor types who were randomized to receive therapy matched to genomic abnormalities and for those who received conventional treatment [47]. The criticisms of this trial reflect the main challenges associated with targeted therapy. A key question is how to identify and classify the variants in genes that encode potential targets.

Advances in sequencing technology have revealed the genomic landscape of multiple cancer types, but the understanding of the functional consequences of these genetic changes has lagged. Being able to discriminate between which variants do, or do not, confer susceptibility to a drug is the key to maximizing response rates. As the number of potential targets increases, then annotating the functional consequences of variants becomes a huge bioinformatic challenge. Robust evidence that matches a drug with a given predictive biomarker is also essential. The SHIVA trial was criticized for not optimizing the match between variant and drug, in part, due to the limited number of targeted drugs currently available.

There is also an increasing appreciation of the influence of tumor heterogeneity on initial response and acquired resistance [48]. Tumor heterogeneity refers to the existence of cell subpopulations with distinct phenotypic characteristics, within (intratumor heterogeneity) and between tumors (intertumor heterogeneity). Cancers evolve from an ancestral cell, this evolution can be fueled by genomic instability and shaped by the selection pressures exerted by treatment. Tumor evolution can be depicted as a branching tree with early truncal mutations being shared by all sites of disease, even in patients with advanced cancers. Mutations arising over time are restricted to sets of subclones, giving rise to a branch-like pattern. For example, BRAF V600 mutations occur in dysplastic nevi before they progress to malignant melanoma and remain as an oncogenic driver during progression to malignant melanoma. Targeting of these truncal mutations might be expected to lead to an initial response. However, the selection pressure exerted may allow resistant subclonal populations to become the dominant cell population, manifesting as treatment failure and disease progression. But even where complex polyclonal resistance emerges, the resistance mechanisms may converge on specific pathways, suggesting this might be managed with drugs that target this pathway [49]. This highlights the importance of increasing the understanding of the evolution of resistance.

Genetic profiling to date has been performed on a tissue sample obtained by a biopsy or resection. However, a tumor's genetic makeup may vary from one part of the tumor to another, from the primary tumor compared to metastatic sites and may change over time. It is not feasible to obtain multiple biopsies from the tumor and all sites of spread nor to repeat biopsies multiple times throughout treatment. Cell-free circulating tumor DNA (ctDNA) consists of DNA fragments released from tumors into the circulation [50]. Multiple cells from within the primary tumor and from metastatic sites contribute to the ctDNA pool, which consequently reflects tumor genetic heterogeneity. As the ctDNA is obtained from blood samples, it is relatively easy to obtain multiple samples, so changes in the amount of ctDNA and the mutational profile can be captured over time. However, ctDNA is present in small amounts in the blood, and most circulating DNA is derived from normal cells. A fundamental challenge in the analysis of ctDNA is that mutated DNA fragments in the blood are present at low concentrations and mutations of interest may be present at very low allele frequency. Detection of mutations at such very low allele frequency requires deep sequencing coverage and error suppression techniques to distinguish true low-level variants from amplification and sequencing errors. Technical advances are allowing the development of strategies that will allow ctDNA analysis to become part of routine clinical practice and changes in a cancers' genetic profile to be tracked during treatment.

---

## 29.6 Making common cancer rare

---

By taking common diseases such as breast or colorectal cancer and stratifying them according to, often individually rare genetic variants, we have converted a small number of common conditions into a large number of individually rare conditions. This has a broad range of implications on areas such as trial design, on the level of the evidence required for regulatory approval and on reimbursement of treatment costs. Large randomized phase III trials may be of decreasing practicality if entry is based on molecularly stratified subtypes rather than broad diagnostic groups. Larger number of potential participants will need to be screened to identify those that are eligible, and treatment centers may need to offer a bigger portfolio of trials to allow the matching of patients with the appropriate experimental targeted drug. Lessons may need to be learned from the rare disease community, including the need for regulators to accept lower quality levels of evidence than that obtained from large phase III trials, which in some cases may be too difficult to conduct.

---

## 29.7 Conclusion

---

Advances in genomics are driving new approaches in many areas of medicine, including the treatment of cancer. Genomics is leading to fundamental insights into the biology of cancer and a new taxonomy of cancer, which

may complement or replace traditional histological classifications systems. This genomic taxonomy may be more informative in terms of predicting prognosis and response and will identify new potential targets. But achieving the full potential of this genetic-based approach requires an increased understanding of the functional consequences of genetic changes, so that therapies can be rationally designed and selected. In the foreseeable future, resistance will remain the key reason why most treatments ultimately fail. Developing strategies to overcome resistance may have as much impact on outcome as further increasing the number and range of targeted drugs.

There have been some spectacular successes for molecularly targeted therapies in oncology. The pace of development is likely to increase further, and drugs for today's intractable conditions will become tomorrow's successes.

## References

- [1] Pagliarini R, Shao W, Sellers WR. Oncogene addiction: pathways of therapeutic response, resistance, and road maps toward a cure. *EMBO Rep* 2015;16(3):280–96.
- [2] Chereda B, Melo JV. Natural course and biology of CML. *Ann Hematol* 2015;94(Suppl 2):S107–21.
- [3] Soverini S, et al. Chronic myeloid leukemia: the paradigm of targeting oncogenic tyrosine kinase signaling and counteracting resistance for successful cancer therapy. *Mol Cancer* 2018;17(1):49.
- [4] Nowell PC, Hungerford DA. Chromosome studies on normal and leukemic human leukocytes. *J Natl Cancer Inst* 1960;25:85–109.
- [5] Rowley JD. Letter: a new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 1973;243(5405):290–3.
- [6] Stagno F, et al. Imatinib mesylate in chronic myeloid leukemia: frontline treatment and long-term outcomes. *Expert Rev Anticancer Ther* 2016;16(3):273–8.
- [7] O'Brien SG, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 2003;348(11):994–1004.
- [8] Hayward NK, et al. Whole-genome landscapes of major melanoma subtypes. *Nature* 2017;545(7653):175–80.
- [9] Kim A, Cohen MS. The discovery of vemurafenib for the treatment of BRAF-mutated metastatic melanoma. *Expert Opin Drug Discov* 2016;11(9):907–16.
- [10] Chapman PB, et al. Vemurafenib in patients with BRAFV600 mutation-positive metastatic melanoma: final overall survival results of the randomized BRIM-3 study. *Ann Oncol* 2017;28(10):2581–7.
- [11] Banzi M, et al. Dabrafenib: a new opportunity for the treatment of BRAF V600-positive melanoma. *Onco Targets Ther* 2016;9:2725–33.
- [12] Gatzemeier U, et al. Phase III study of erlotinib in combination with cisplatin and gemcitabine in advanced non-small-cell lung cancer: the Tarceva Lung Cancer Investigation trial. *J Clin Oncol* 2007;25(12):1545–52.
- [13] Lynch TJ, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004;350(21):2129–39.
- [14] Inoue A, et al. Prospective phase II study of gefitinib for chemotherapy-naïve patients with advanced non-small-cell lung cancer with epidermal growth factor receptor gene mutations. *J Clin Oncol* 2006;24(21):3340–6.
- [15] Mok TS, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;361(10):947–57.
- [16] Spagnuolo A, Maione P, Gridelli C. Evolution in the treatment landscape of non-small cell lung cancer with ALK gene alterations: from the first- to third-generation of ALK inhibitors. *Expert Opin Emerg Drugs* 2018;23(3):231–41.
- [17] Lin JJ, Shaw AT. Recent advances in targeting ROS1 in lung cancer. *J Thorac Oncol* 2017;12(11):1611–25.
- [18] Bivona TG, Doebele RC. A framework for understanding and targeting residual disease in oncogene-driven solid cancers. *Nat Med* 2016;22(5):472–8.
- [19] Sabnis AJ, Bivona TG. Principles of resistance to targeted cancer therapy: lessons from basic and translational cancer biology. *Trends Mol Med* 2019;25(3):185–97.
- [20] Murtuza A, et al. Novel third-generation EGFR tyrosine kinase inhibitors and strategies to overcome therapeutic resistance in lung cancer. *Cancer Res* 2019;79(4):689–98.
- [21] Kobayashi S, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2005;352(8):786–92.
- [22] Soria JC, et al. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. *N Engl J Med* 2018;378(2):113–25.
- [23] Flaherty KT, et al. Inhibition of mutated, activated BRAF in metastatic melanoma. *N Engl J Med* 2010;363(9):809–19.
- [24] Nazarian R, et al. Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature* 2010;468(7326):973–7.
- [25] Turajlic S, et al. Whole-genome sequencing reveals complex mechanisms of intrinsic resistance to BRAF inhibition. *Ann Oncol* 2014;25(5):959–67.
- [26] Nissan MH, et al. Loss of NF1 in cutaneous melanoma is associated with RAS activation and MEK dependence. *Cancer Res* 2014;74(8):2340–50.
- [27] Lim SY, Menzies AM, Rizos H. Mechanisms and strategies to overcome resistance to molecularly targeted therapy for melanoma. *Cancer* 2017;123(S11):2118–29.
- [28] Sinha R, et al. Cutaneous toxicities associated with vemurafenib therapy in 107 patients with BRAF V600E mutation-positive metastatic melanoma, including recognition and management of rare presentations. *Br J Dermatol* 2015;173(4):1024–31.
- [29] Robert C, et al. Improved overall survival in melanoma with combined dabrafenib and trametinib. *N Engl J Med* 2015;372(1):30–9.
- [30] Pazarentzos E, Bivona TG. Adaptive stress signaling in targeted cancer therapy resistance. *Oncogene* 2015;34(45):5599–606.
- [31] Bivona TG, et al. FAS and NF- $\kappa$ B signalling modulate dependence of lung cancers on mutant EGFR. *Nature* 2011;471(7339):523–6.
- [32] Brunen D, Bernards R. Drug therapy: exploiting synthetic lethality to improve cancer therapy. *Nat Rev Clin Oncol* 2017;14(6):331–2.

- [33] Farmer H, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* 2005;434(7035):917–21.
- [34] Palazzo L, Ahel I. PARPs in genome stability and signal transduction: implications for cancer therapy. *Biochem Soc Trans* 2018;46(6):1681–95.
- [35] Chen CC, et al. Homology-directed repair and the role of BRCA1, BRCA2, and related proteins in genome integrity and cancer. *Annu Rev Cancer Biol* 2018;2:313–36.
- [36] Mazières J, et al. Evaluation of EGFR protein expression by immunohistochemistry using H-score and the magnification rule: re-analysis of the SATURN study. *Lung Cancer* 2013;82(2):231–7.
- [37] Lord CJ, Ashworth A. PARP inhibitors: synthetic lethality in the clinic. *Science* 2017;355(6330):1152–8.
- [38] Bitler BG, et al. PARP inhibitors: clinical utility and possibilities of overcoming resistance. *Gynecol Oncol* 2017;147(3):695–704.
- [39] Davies H, et al. HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat Med* 2017;23(4):517–25.
- [40] Nebot-Bral L, et al. Hypermutated tumours in the era of immunotherapy: the paradigm of personalised medicine. *Eur J Cancer* 2017;84:290–303.
- [41] Baretto M, Le DT. DNA mismatch repair in cancer. *Pharmacol Ther* 2018;189:45–62.
- [42] Battaglin F, et al. Microsatellite instability in colorectal cancer: overview of its clinical significance and novel perspectives. *Clin Adv Hematol Oncol* 2018;16(11):735–45.
- [43] Yan L, Zhang W. Precision medicine becomes reality-tumor type-agnostic therapy. *Cancer Commun (Lond)* 2018;38(1):6.
- [44] Kheder ES, Hong DS. Emerging targeted therapy for tumors with NTRK fusion proteins. *Clin Cancer Res* 2018;24(23):5807–14.
- [45] Drlon A, et al. Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children. *N Engl J Med* 2018;378(8):731–9.
- [46] Hyman DM, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. *N Engl J Med* 2015;373(8):726–36.
- [47] Le Tourneau C, et al. Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. *Lancet Oncol* 2015;16(13):1324–34.
- [48] McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* 2017;168(4):613–28.
- [49] Venkatesan S, Birkbak NJ, Swanton C. Constraints in cancer evolution. *Biochem Soc Trans* 2017;45(1):1–13.
- [50] Corcoran RB, Chabner BA. Cell-free DNA analysis in cancer. *N Engl J Med* 2019;380(5):501–2.





# Gene, genome, and molecular therapeutics

*Dhavendra Kumar*<sup>1,2,3</sup>

<sup>1</sup>Institute of Medical Genetics, University Hospital of Wales, Cardiff, United Kingdom <sup>2</sup>Genomic Policy Unit, University of South Wales, Pontypridd, United Kingdom <sup>3</sup>William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

## 30.1 Introduction

Apart from understanding cellular and molecular mechanisms, the strength of molecular medicine is largely appreciated by diagnostic precision and exploiting molecular pathology for specific therapeutic interventions. In the previous chapters, information and discussion are confined to genetic and genomic basis of drug metabolism and drug response. In the current chapter, recombinant drugs and new therapeutic avenues targeting at specific genes, molecules, and cellular structures are discussed. There are many such approaches largely confined to biomedical research. However, few are now approved for clinical use, and the list is fast expanding. The chapter includes separate sections on stem-cell therapy, gene therapy, antisense oligonucleotides (AS ONs), ribozymes, RNA interference (RNAi), aptamers, and CRISPR/Cas9 gene/genome editing. In keeping with the size and scope of the book, scientific and experimental details are avoided. Emphasis is given to potential clinical applications. There are key references cited at the end of chapter to assist interested reader and student for further reading.

## 30.2 Recombinant protein drugs and vaccines

The invention and large-scale use of recombinant protein drugs are by far the best example of targeted molecular therapeutics. Insulin was the first recombinant protein drug that was approved by the United States Food and Drug Agency (<https://www.fda.gov/>) in 1982, followed by around 250 drugs produced by heterologous expression. This area is perhaps the biggest revenue earner for the pharmaceutical industry estimated to be approximately \$150 billion. A number of diseases including cancer, diabetes mellitus, cardiovascular disease, inflammatory disease, and microbial and genetic diseases are now successfully managed by different classes of recombinant drugs. Some of these agents have revolutionized the treatment for a number of rare diseases, commonly referred to orphan drugs. Most importantly, some of these new classes of drugs are used for common diseases, such as diabetes mellitus [1].

### 30.2.1 Recombinant pharmacotherapy

In early 1980s the development of recombinant DNA technologies provided a path for the production of protein drugs in expression organisms, notably *Escherichia coli*. This procedure allowed large-scale production of markedly pure and safe proteins that were isolated from animal or human sources. In addition, the pharmacological properties of these proteins can be altered by the introduction of mutations or posttranslational modifications, for example, the attachment of covalent polyethylene glycol (PEG). Following the human genome sequencing, novel human proteins that were previously not available can now be designed and synthesized. Large-scale protein production is possible using *E. coli* cultivated in huge quantity in bioreactor microorganisms.

**TABLE 30.1** Classes of recombinant protein drugs.

Class of recombinant drug	Examples
Monoclonal antibodies	Anti-Her2, anti-TNF, anti-VEGF
Hormones	Insulin, glucagon-like peptide-1, growth hormone
Growth factors	Erythropoietin, granulocyte colony-stimulating factor
Fusion proteins	TNF-receptor-Fc
Cytokines	Interferon- $\alpha$ , interferon- $\beta$ , interferon- $\gamma$
Blood coagulation factors	Factor VIII, VIIa, IX
Anticoagulants and thrombolytics	Tissue-type-plasminogen activator
Therapeutic enzymes	Glucocerebrosidase
Vaccines	Human papilloma virus L1 protein Hepatitis B surface antigen

TNF, tumour necrosis factor; VEGF, vascular endothelium growth factor.

Adapted with permission from Kurreck J, Stein CA. *Molecular medicine—an introduction*. Wiley-VCH., 2016, p. 239, Table 10-4.

**TABLE 30.2** Protein therapeutics—PEGylation: posttranslational bioengineering of protein biotherapeutics.

Brand	Active substance	Indication	Approval year
Adagen	Adenosine deaminase	Severe combined immunodeficiency disease (SCID)	1990
Oncaspar	Asparaginase	Leukemia	1994
Neulasta	Granulocyte-colony stimulating factor (G-CSF)	Neutropenia	2002
PegIntron	Interferon $\alpha$ 2b	Hepatitis C	2000
PEGASYS	Interferon $\alpha$ 2a	Hepatitis C	2002
Mircera	Erythropoietin (EPO)	Anaemia associated with chronic kidney disease	2007

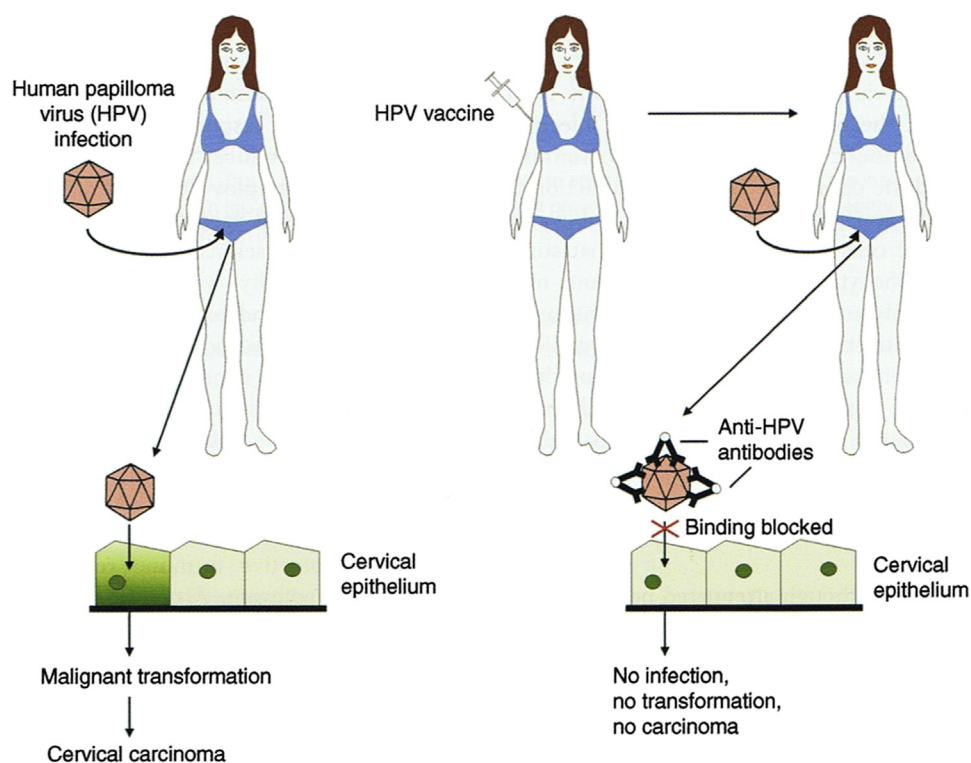
Veronese FM, Pasut G. PEGylation: posttranslational bioengineering of protein biotherapeutics. *Drug Discov Today Technol* 2008;5 (2–3):e57–64 [9].

However, more complex proteins, for example, glycosylated proteins, require more demanding eukaryotic cell expression systems. The complex biotechnological details of the production of recombinant protein drugs are beyond the scope of this chapter.

Currently, different classes of recombinant drugs are used in clinical practice (Table 30.1). The major categories include monoclonal antibodies, hormones, growth factors, fusion proteins, cytokines, blood coagulation factors and thrombolytics, therapeutic enzymes, and recombinant vaccines. In addition, new classes of recombinant protein drugs are produced by posttranslational chemical modification. This method allows the production of drugs with substantially improved pharmacokinetic and pharmacodynamics properties. The attachment of the PEG or alanine and serine (PAS) increases the size of the protein molecule. This approach allows enhanced bioavailability by reducing renal clearance. A number of PEGylated proteins and peptides are now used in clinical practice (Table 30.2).

### 30.2.2 Recombinant vaccines

Recombinant DNA technology has made it possible to produce specific viral antigens for vaccination either acting on the viral surface or capsid proteins. In contrast to attenuated or inactivated microorganisms the entire pathogen is not needed, either for the production of vaccine or for the vaccination. This approach is safe and does not cause the vaccination-induced infection. An important limiting factor in this approach is reduced immunogenic potential since the intact virus particle (virion) is not used. However, the use of an adjuvant might help in enhancing the immunogenicity.



**FIGURE 30.1** The HPV vaccine. HPV vaccine—the anti-HPV subunit vaccine induces an antibody-based immune response that prevents the virus from infecting the epithelial cells and protects against subsequent tumor development. HPV, Human papilloma virus. Source: Adapted with permission from Kurreck J, Stein CA. *Molecular medicine—an introduction*. Wiley-VCH, 2016, p. 138, Figure 5.12.

### 30.2.2.1 Human papilloma virus vaccine

The DNA recombinant human papilloma virus (HPV) vaccine is a major development for the prevention of genital warts in both males and females and cervical cancer in young females. HPV types 6 and 11 are implicated in genital warts, while HPV 16 and HPV 18 are causally linked with cervical cancer. Recombinant HPV vaccines contain the major capsid HPV protein and L1, which can spontaneously self-assemble into virus like particles (VLPs) that resemble HPV. VLPs trigger antibody production that confer protection against HPV infection but do not induce cancer since they lack the viral DNA (Fig. 30.1) [3].

HPV vaccination is nearly effective in 100% cases by preventing cancerous cellular changes. Currently, Cervarix and Gardasil HPV vaccines are available consisting of the L1 protein of HPV 16 and HPV 18. The Gardasil has the advantage of preventing the cervical cancer but also the genital warts, specifically caused by HPV 6 and HPV 11.

### 30.2.2.2 The Hepatitis B recombinant vaccine

The recombinant hepatitis B vaccine (Recombivax and Engerix-B, etc.) has been used since 1980s. It replaced the previous class of vaccines developed by extracting viral surface protein from the blood of chronically infected patients. The subunit (HBsAG) contains the hepatitis B surface antigen, one of the viral envelope proteins. It is effective in 80%–85% cases. An adjuvant can be used for enhancing the immunogenicity.

### 30.2.2.3 HIV vaccines

Diagnosis, treatment, and prevention of the HIV-1 infection are huge challenges faced by clinicians, microbial, and biomedical scientists. The outcome of most clinical trials for anti-HIV-1 vaccines has been disappointing. This is largely due to genetic diversity and mutability of HIV-1 virus. The clinical trial, known as RV144, demonstrated modest reduction of HIV-1 infection in the vaccinated group. However, it was not enough for clinical use. In this vaccine, two HIV vaccines (ALVAC-HIV-1 and AIDSVAX B/E) were combined that were individually ineffective. The ALVAC-HIV-1 is a live viral vector that produces three HIV-1 genes (env, gag, and pol). The genetically engineered protein gp120 boosted the immune effect of the AIDSVAX B/E. A very recent and

innovative approach is the direct delivery of neutralizing antibodies against the virus by gene transfer rather than stimulating the immune system. Initial trials are promising but not yet shown to be effective.

#### 30.2.2.4 DNA vaccines

The development of DNA vaccines is relatively new approach that is a special form of gene therapy. It is based on creating a plasmid to produce an antigen of the pathogen. The DNA vaccine is injected into cells either by syringe or a gene gun. The cellular machinery then produces the viral proteins that are displayed on the cell surface and triggers the immune response. There is no risk of DNA vaccine–induced infections. Since DNA is a very stable molecule, it can be easily handled and distributed. A major advantage of DNA vaccines is the ability to adapt to new viruses or new variants (mutants) of a known virus by simply altering the DNA sequence of the expression plasmid. Despite promising biotechnological superiority, DNA vaccines are not shown to be efficient for clinical use.

### 30.2.3 Recombinant therapeutic enzymes

In recent years the management of rare inherited metabolic storage diseases has significantly improved with the use of recombinant therapeutic enzymes. Essentially, the enzyme-replacement therapy (ERT) is centered on the concept of cellular replacement of missing or inadequate enzyme essential for the biodegradation of lysosomal proteins leading to progressive accumulation of biological waste product in varying amounts in different tissues and organs. Most disorders are rare or ultrarare and encountered infrequently in clinical practice. These are also referred to as rare diseases. Technically, the production of recombinant therapeutic enzymes is enormously expensive. The annual cost for treating an average patient could be approximately US \$200,000.

Currently, several therapeutic enzymes are commercially available (Table 30.3). While the outcome of early enzyme replacement is clinically encouraging, overall long-term prognosis is not too promising. Nevertheless, this approach is more acceptable to reduce or mitigate life-threatening complications or probably improved neurodevelopmental and cognitive performance. Most inherited metabolic tertiary units use ERT in lysosomal storage diseases, particularly, Gaucher disease (Cerezyme and Vpriv), Fabry's syndrome (Fabrazyme and Replagal), and X-linked Hunter's disease (Elaprase). Pulmozyme is a recombinant deoxyribonuclease used for treating cystic fibrosis.

## 30.3 Stem-cell therapy

The tissue-specific stem-cell transplantation is now clinically used for the treatment of hematopoietic malignancies. Essentially, it is an example of specific tissue or cellular genome transplantation. The focus of biomedical stem cell research is currently on clinical applications of pluripotent stem cells (PSCs), derived from either human embryos stem cells (hESCs) or induced PSCs (iPSCs) derived from healthy individuals of patients affected with a

**TABLE 30.3** Enzyme-replacement therapy in rare inherited metabolic disorders.

Rare genetic disease	Enzyme involved	Therapeutic enzyme
Gaucher disease (types 1, 2, and 3)	$\beta$ -Glucocerebrosidase	Cerezyme, Vpriv, Elelyso
Hunter syndrome (MPSII)	Iduronate-2-sulfatase	Elaprase (idursulfase)
Fabry disease	Alpha-galactosidase A	Fabrazyme, Replagal (Alpha-Gal-A)
Pompe disease	Alpha-glucosidase (GAA)	Lumizyme (Genzyme)
Maroteux–Lamy syndrome (MPSVI)	Arylsulfatase B (ARSB)	Naglazyme (galsulfase)
Morquio A syndrome	N-Acetylgalactosamine-6-sulfatase (GALNS)	Vimizim (elosulfase alfa)
LAL deficiency	LAL	Kanuma (sebelipase)
Cystic fibrosis	CFTR	Pump zyme

CFTR, Cystic fibrosis transmembrane conductance regulator; LAL, lysosomal acid lipase.

Adapted from [www.raredr.com](http://www.raredr.com).

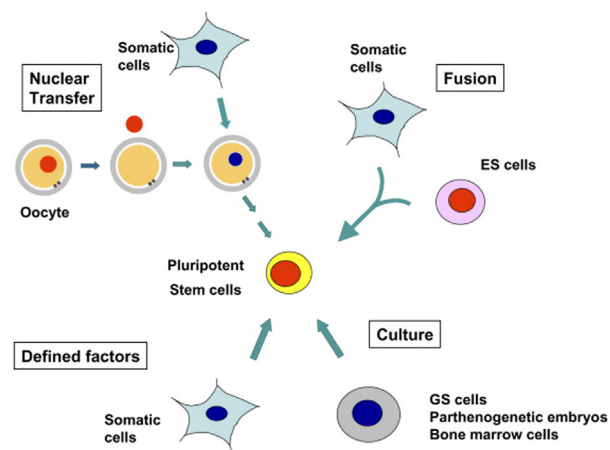
known genetic disease mutation. Initial drive for stem-cell therapy was focused on the use of hESCs. However, due to technical difficulties on harvesting good amount of stem cells and ensuing ethical and moral challenges led to the search for alternative methods, particularly iPSCs.

There are serious ethical considerations that need to be addressed, particularly the use of hESCs. It is imperative that an embryo would need to be destroyed for harvesting the stem cells. To some extent, this could be managed for biomedical research since the selected wild embryo was close to its storage limit (normally 5 years) or going to be discarded for being over the embryo numbers (normally up to three) allowed for in vitro fertilization. Usually, there is no ethical objection for the embryo with a specific gene mutation or pathogenic variant. However, in clinical practice, the situation would be different. It is important to resolve this dispute through serious social–ethical debate supported by relevant national statutory framework.

iPSCs are a special type of PSCs and are obtained by reprogramming differentiated adult cells. These cells have the capacity to propagate indefinitely or can be differentiated into other cell types in the body, for example, into neurons, cardiomyocytes, pancreatic, and liver cells. They are similar to hESCs but have some important advantages. Since these are produced directly from somatic cells, they bypass the need to destroy embryos, eliminating ethical considerations. Moreover, iPSCs can be regenerated in a patient-specific environment, allowing the generation of autologous transplants without the risk of immune rejection. The iPSCs-based stem cell therapy holds great promise in the field of regenerative medicine. Prof. Shinya Yamanaka in Kyoto, Japan pioneered the iPSC technology, who successfully demonstrated the introduction of four specific genes converted differentiated adult fibroblast cells in iPSCs (Fig. 30.2). The 2012 Nobel Prize for Physiology and Medicine was awarded to Yamanaka along with Sir John Gurdon.

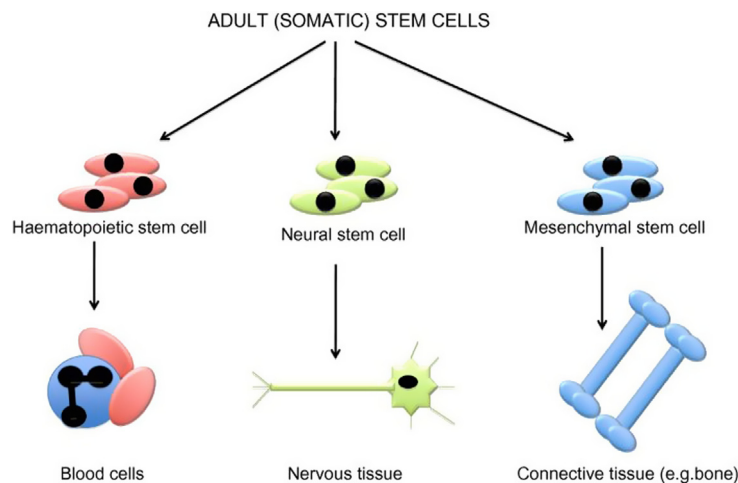
Currently, the focus is on harvesting iPSCs and widespread use in different clinical conditions. Techniques for harvesting PSCs into any of the three germinal layers (ectoderm, mesoderm, and endoderm) are available and routinely used in new drug development and specific tissue banking [4]. The most important approach is focused on patient-specific iPSCs for the development of a cell type used as a disease model. The differentiated cell type can be used to study disease mechanisms or to screen libraries for active compounds that eventually can be used for treatment (Fig. 30.3).

Another therapeutic model of PSCs is whole cell therapy in a number of chronic diseases. The logic underpinning this approach is anticipated conversion of transplanted PSCs in the required cell type and in appropriate natural growth environment of naturally occurring growth enhancers and growth factors. Clinical examples include amyotrophic lateral sclerosis (motor neurone disease), spinal cord injury, ischemic heart disease (myocardial infarction), type 2 diabetes mellitus, and ischemic stroke. The list of conditions managed by this approach is continually expanding (Fig. 30.4). Details on PSCs therapeutics in individual diseases are outside the scope of this chapter. In this context, it is important to consider few generic issues and challenges for medical applications of PSCs.

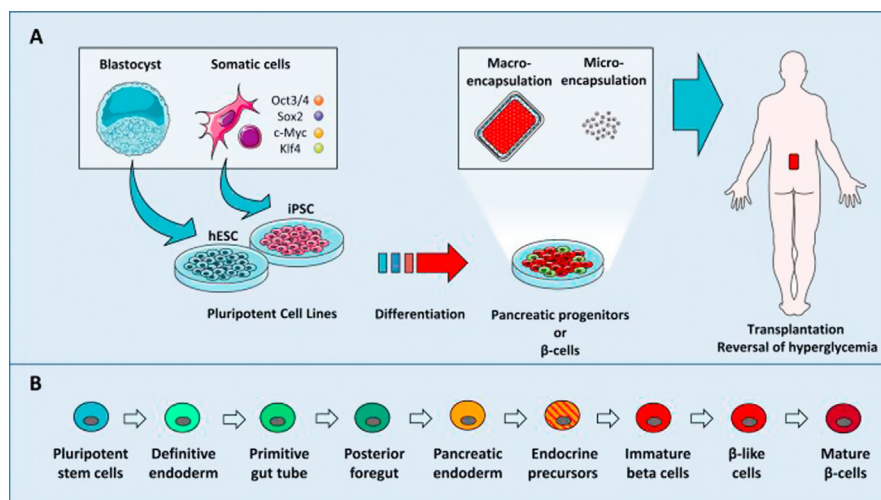


**FIGURE 30.2** The Yamanaka model of generation of iPSCs by reprogramming the adult fibroblasts using four transcription factors. Source: Adapted with permission from Yamanaka, S, *Strategies and New Developments in the Generation of Patient-Specific Pluripotent Stem Cells*. *Cell Stem Cell* 2007; 1(1):39–49 [5].





**FIGURE 30.3** Medical usage of iPSCs. Source: Adapted with permission from Gattegno-Ho, D, Argyle, S, Argyle, DJ, *Stem cells and veterinary medicine: Tools to understand diseases and enable tissue regeneration and drug discovery. The Veterinary Journal* 2012;191(1):19–27 [6].



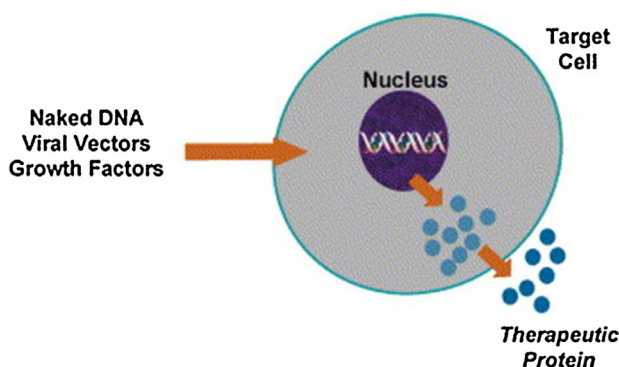
**FIGURE 30.4** Potential use of cells differentiated from pluripotent stem cells. Source: Adapted with permission from Quiskamp, N, Bruin, JE, Kieffer, TJ, *Differentiation of human pluripotent stem cells into  $\beta$ -cells: Potential and challenges. Best Practice & Research Clinical Endocrinology & Metabolism* 2015;29(6):833–47 [7].

Unwanted side effects of PSCs remain a major issue. The risk for teratoma formation by PSCs injection is a known hazard, comparatively more with iPSCs than hEPSCs. Thus direct transplantation of PSCs could not be justified and should be prohibited. Techniques for in vitro differentiation would need to be developed and validated for transplantation of specific differentiated PSCs. However, the risk of teratoma would still be there but probably clinically justified similar to risks associated with any other therapeutic interventions. New techniques of transcellular differentiation might offer some safety net of the potential neoplasia risks, for example, reprogramming the harvested fibroblasts developed from iPSCs. Another approach for reducing the risk is to deliver programmed PSCs directly in the targeted tissue or region rather than through direct infusion in the blood stream.

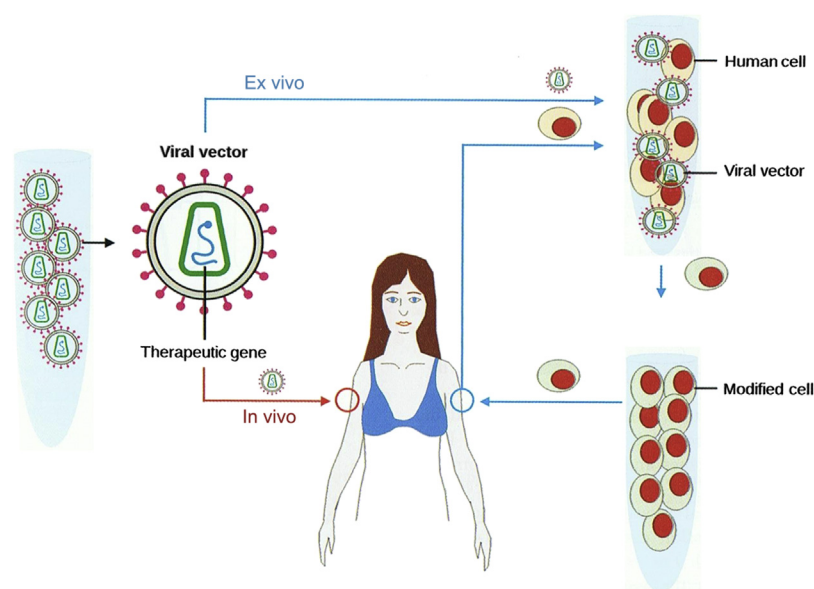
The success of iPSCs led cellular regeneration is largely confined to nongenetic causes, for example, ischemia in myocardial infarction and stroke, spinal injury, or autoimmune-mediated pancreatic cellular loss in type 1 diabetes mellitus. However, this approach cannot cure some diseases caused by a genetic defect. Currently, the focus is now on combined strategies of gene therapy and iPSCs stem-cell therapy. This is discussed in the following section.

## 30.4 Gene therapy

Gene therapy is one of the much desired and anticipated therapeutic paradigms of clinical molecular medicine. It is essentially a molecular method involving the transference of normal coding genetic material into the



**FIGURE 30.5** Principles of gene therapy. Source: Adapted with permission from Ibraheem, D, Elaissari, A, Fessi, H, *Gene therapy and DNA delivery systems*. International Journal of Pharmaceutics, 2014;459(1–2), 70–83 [8].



**FIGURE 30.6** Ex vivo and in vivo gene therapy. Source: Adapted with permission from Kurreck J, Stein CA. *Molecular medicine—an introduction*. Wiley-VCH, 2016, p. 259, Figure 11.2.

cells of a patient (Fig. 30.5). The gene transfer could be either episomal (not integrated in the host genome) or integrated in patient's cells. The defective gene remains intact as before. Earlier gene therapy strategies involved monogenic diseases. Now this approach is also applied to cancer, cardiovascular disease, infectious diseases, and many other disorders.

The fundamental aspect of gene therapy is the selection of vector to be used as a vehicle for gene (selected wild sequences) transfer. Most of these are viral vectors, particularly retrovirus, adenovirus, and adeno-associated viruses. Other vectors have also been tried, for example, liposomes in cystic fibrosis. The use of viral vectors is not without potential risks and side effects, for example, insertional mutagenesis, local neoplastic proliferation, and unwanted immunogenicity. In addition, vectors also differ in duration of action, tissue distribution, and efficiency. Another inhibiting factor for viral vectors is the lack of replication. They can only induce patient's cells but without any power for dissemination.

Techniques of gene transfer using any form of vector are highly complex and are not discussed here. However, broadly there are two types of gene therapy transfer system—germ line and somatic. In germ line the gene transfer is aimed in the ovum or sperm. Unfortunately, this approach led to serious ethical objections, excessive transmission of new genetic risks or side effects in any future offspring of children conceived with germ line therapy. Whilst the germ line gene therapy remains elusive, this modality is not favored and has no clinical applications. In contrast the somatic gene therapy is preferred since the genetic change is restricted to the patient and is not passed on to his or her children. The somatic gene therapy is either ex vivo or in vivo (Fig. 30.6). In ex vivo method, cells are removed from the patient (e.g., bone marrow biopsy) and then genetically modified

**TABLE 30.4** Gene therapy applications in monogenic diseases.

Disorder (OMIM number <sup>a</sup> )	Disorder type	Mutated gene	Function	Inheritance
Cystic fibrosis (219700)	Channelopathy	CFTR	Chloride channel	Autosomal recessive
Duchenne muscular dystrophy (310200)	Musculoskeletal	Dystrophin	Cytoskeletal component	X-linked
Fabry disease (301500)	Metabolic	$\alpha$ -Galactosidase A	Glycosphingolipid catabolism	X-linked
Huntington's disease (143100)	Musculoskeletal	Huntingtin	Vesicular trafficking mediator <sup>b</sup>	Autosomal dominant
Sickle cell anaemia (603903)	Haematological	$\beta$ -Globin	Oxygen carrier	Autosomal recessive
X-linked severe combined immunodeficiency (300400)	Immune dysfunction	Interleukin-2 receptor gamma subunit	Cytokine receptor	X-linked

Wong, G.K., Chiu, A.T., *Gene therapy, gene targeting and induced pluripotent stem cells: applications in monogenic disease treatment. Biotechnology Advances* 2011;29 (1):1–10.

with the vector in tissue culture before being implanted. In contrast the *in vivo* gene therapy involves treating patient directly with the vector. In both types the therapeutic benefit could be of limited duration or permanent.

Since 1990s, more than 2000 clinical trials of gene therapy have been undertaken costing millions of dollars. Unfortunately, limited numbers of monogenic conditions are approved for gene therapy treatment strategy due to serious problems of efficacy and side effects (Table 30.4). However, biomedical research in gene therapy continues exploring and discovering newer classes of vectors.

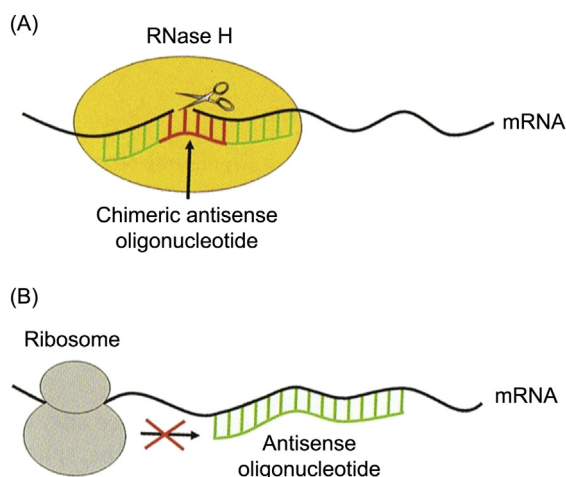
### 30.5 Antisense oligonucleotides

The fundamental basis of all anti-mRNA strategies is the oligonucleotides that specifically bind to a target mRNA by Watson–Crick base pairing and prevent its translation into a protein. These techniques allow studying the function of a gene by producing a “loss of function” phenotype or for treating a disease by inhibiting the expression of a deleterious gene. This section describes some of the basic concepts and clinical applications of AS ONs.

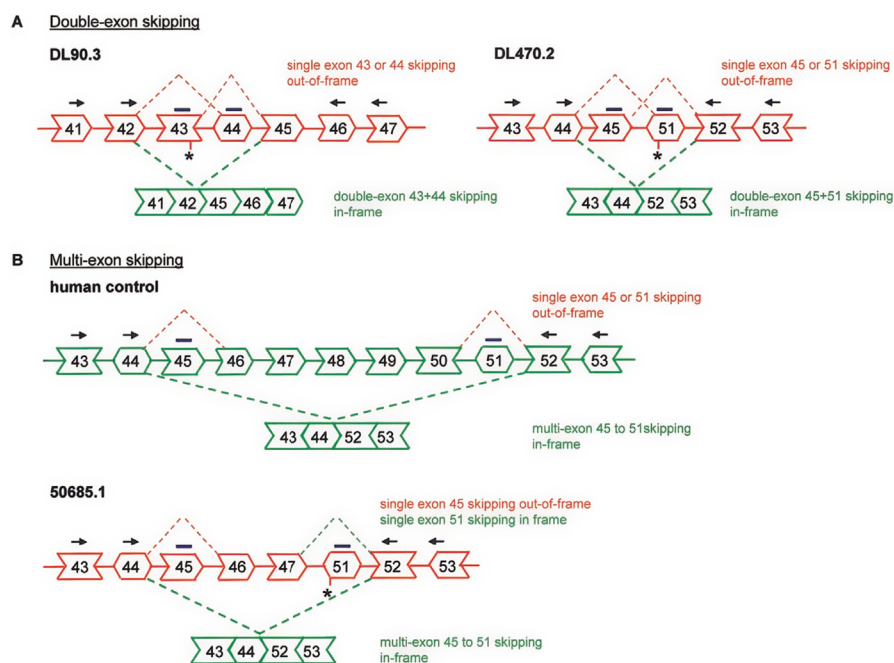
Most AS ONs are usually 15–20 nucleotides (nt) long and inhibit translation by blocking ribosomal translation of mRNA into protein or by inducing cleavage of the target RNA by RNase H (Fig. 30.7) [11]. The first clinically relevant use of AS ONs was inhibiting the Rous Sarcoma virus. However, despite technical advances, development of the therapeutically useful AS ONs faces numerous challenges. Most important hurdles include the identification of specific regions in the target mRNA, the stabilization of the AS ONs against nucleolytic degradation, and the efficient delivery of the AS ONs into cells of the target tissue. Complex laboratory processes allow to overcome technical hurdles that are beyond the scope of this small section.

So far around 30 AS ONs have been tested in different clinical trials leading to the approval for therapeutic use. Since the antisense technology can be applied to potentially inhibit the expression of any gene, its clinical applications range from oncology to viral infection. Most of the AS ONs clinical trials are based on phosphorothioates, now approved by the United States Federal Drug Agency (FDA) for the treatment of cytomegalovirus (CMV) retinitis in HIV-infected patients, a leading cause of blindness. The drug (Fomivirsen), marketed as Vitravene, is administered by intravitreal route and inhibits CMV and can prevent blindness. Another FDA-approved drug (Mipomersen/Kynamro) is used for homozygous hypercholesterolemia by inhibiting apolipoprotein B.

AS ON therapy can also be used without the degradation of target mRNA. This approach is used in a novel therapeutic success for the treatment of severe form of X-linked Duchenne muscular dystrophy (DMD) by converting to the mild form known as Becker muscular dystrophy. The basis of this therapy is exon skipping along with adjoining introns. This allows series of exons continuously lined up for transcription, and thus protein production eventually, however, decreases both in quantity and quality (Fig. 30.8). It has significant limitation since in frame exonic dystrophic deletion is only encountered in less than 15% of DMD patients. However, work continues to create personalized AS ONs for specific dystrophin point or missense mutations.



**FIGURE 30.7** Antisense oligonucleotides consist of DNA or modified DNA analogs that bind to target RNA to inhibit gene expression. (A) RNase H cleavage and (B) blocking of translation. Source: Adapted with permission from Kurreck J, Stein CA. *Molecular medicine—an introduction*. Wiley-VCH, 2016, p. 305, Figure 13.3.



**FIGURE 30.8** Antisense oligonucleotides for exon skipping strategy in Duchenne muscular dystrophy. Source: Adapted with permission from Aartsma-Rus, A, Janson, AA, Kaman, WE, Bremmer-Bout, M, van Ommen, GJ, den Dunnen, JT, van Deutekom, JC. *Antisense-Induced Multiexon Skipping for Duchenne Muscular Dystrophy Makes More Sense*. *Am J Hum Genet*. 2004;74 (1):83–92 [10].

Despite AS ONs being clinically well tolerated with fewer toxic side effects, however, the efficiency is not too high to allow clinical approval. This led to the development of other RNA inhibition models discussed in latter sections.

## 30.6 Ribozymes

Ribozymes are unique RNA molecules located in ribosomes and possess enzymatic activity and as well as act as carriers for genetic information. In the early 1980s, Noble Laureates Thomas Cech and Sydney Altman discovered ribozymes by demonstrating that the RNA of RNase P is active without a protein component. Previously the enzymatic activity was attributed to proteins, while nucleic acids were regarded as simply carriers of genetic information. These findings led to the ensuing debate on the existence of “RNA world hypothesis.”

There are two main classes of ribozymes—large, comprising hundreds and thousands of nucleotides, and small, made of around 30–150 nt. In addition, there are few others with that possess ribozymes like catalytic activity [12]. Detailed discussion on all forms of ribozymes is outside the scope of this small section. For medical purposes the main function of ribozymes exploited includes degradation or correction of mutated target mRNA involved in the  $\beta$ -globin peptide chain assembly. The splicing activity of group I introns is used to replace pathologically mutated section of the  $\beta$ -globin mRNA with the intact exon. Group I introns interact with the 5' exon via an internal guide sequence (IGS). An artificial ribozyme is used made up from complementary sequences to target mRNA replacing the natural IGS. The downstream exon segment of the group I intron is also modified to contain the intact version of the  $\beta$ -globin mRNA instead of the natural sequence. This approach succeeded in correcting the defect and no longer led to pathologic hemoglobin polymerization.

So far, several hundreds of ribozymes have been evaluated for clinical applications [13]. Among these, clinically important classes of ribozymes are hammerhead ribozymes, chemically synthesized ribozymes, and intracellularly expressed ribozymes after vector transfer. So far the main focus is for targeting certain viruses and tumor angiogenesis. Clinical trials using Angiozyme, an artificially produced chemically synthetic ribozyme, have shown promising results in blocking tumor angiogenesis. It is directed against the vascular endothelial growth factor. Angiozyme inhibited tumor growth in several mouse models. Unfortunately, clinical trials were stopped in phase 2 after failing to achieve clinically effective outcomes.

Other clinical use is designed to target the human epidermal growth factor type 2 (Her2) that is overexpressed in breast cancer with widespread metastasis and poor prognosis. Clinical trials involving Herzyme, a class of hammerhead ribozyme, have provided encouraging early results. Another chemically synthesized ribozymes is being evaluated against hepatitis C. The antiviral activity of ribozymes is enhanced with vectors. This is clinically evaluated for treating HIV-1 infection. The ribozyme, OZI, is introduced against the overlapping *vpr* and *tat* reading frames of HIV-1. A phase II clinical trial proved safe and effective in HIV-infected patients.

### 30.7 RNA interference

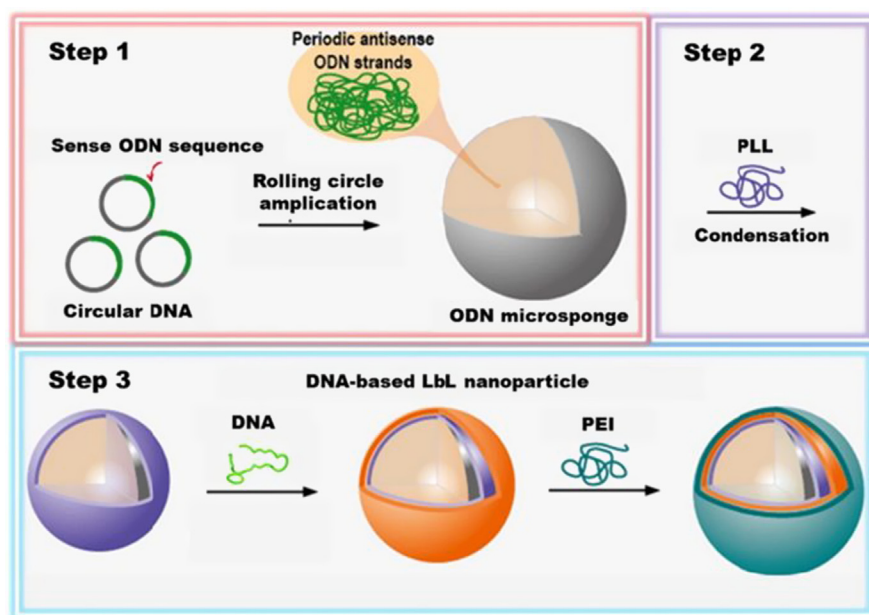
Since the advent of RNAi in 1990s, it has become one of the most preferred laboratory methods for posttranscriptional gene silencing [14]. The 2006 Nobel Prize for Medicine and Physiology to Andrew Fire and Craig Mello recognizes this field of research. Practically RNAi was shown to silence any gene and has been the subject of many high profile molecular biology research projects. The therapeutic potential of RNAi is considered to be far more robust compared to AS ONs and ribozymes.

Molecular mechanisms involved in the gene silencing by RNAi are very complex. An important component of the process includes cutting down of the double stranded RNA to short or small interfering RNA (siRNA). The siRNAs are engaged in gene silencing through RNA-induced silencing complex (RISC) with the aide of RISC-loading complex. It is now established that siRNAs specifically and efficiently inhibit the expression of a target gene in mammalian cells [15]. Since RNAi occurs in the cytoplasm, it has the added advantage of direct involvement in the process of RNA cleavage. In contrast, the AS ON goes through complex nuclear pathway (Fig. 30.9). Details on many other complex cellular and molecular aspects of RNAi in gene silencing are beyond the scope of this small section.

A number of preclinical “proof of principle” studies provided sufficient evidence of RNAi in promising clinical applications. There are now large numbers of RNAi therapeutic clinical trials under evaluation. Notable examples include HIV infection, hepatitis B viral infection, metastatic melanoma, hepatic cancer, hypercholesterolemia, ovarian cancer, colorectal cancer, and glaucoma. An important disease likely to benefit is the age-related macular degeneration (AMD). The wet type AMD is clinically similar to edematous diabetic retinopathy. The RNAi-based drug Bevasiranib is administered locally by retinal injection. It inhibits the vascular proliferation by silencing VEGF gene or its receptor. However, later studies indicated that actual inhibition of the vascular proliferation was achieved by activating the Toll-like receptor 3 (TLR3) [17].

In addition to eye disease, clinical applications of RNAi gene-silencing technology are under way in many other diseases including cancer and viral infection. One of the most commercially exciting developments of RNAi is evident from permanent hair removal. This is achieved by silencing the *Hairless* gene in the hair follicle [18]. This application is estimated to be worth \$10 billion.





**FIGURE 30.9** Diagrammatic illustration showing comparison of antisense oligonucleotides and RNA-interference technologies. Source: Adapted with permission from Jeong, E, Kim, H, Jang, B, Cho, H, Ryu, J, et al., *Technological development of structural DNA/RNA-based RNAi systems and their applications*. *Advanced Drug Delivery Reviews*, 2016;104:29–43 [16].

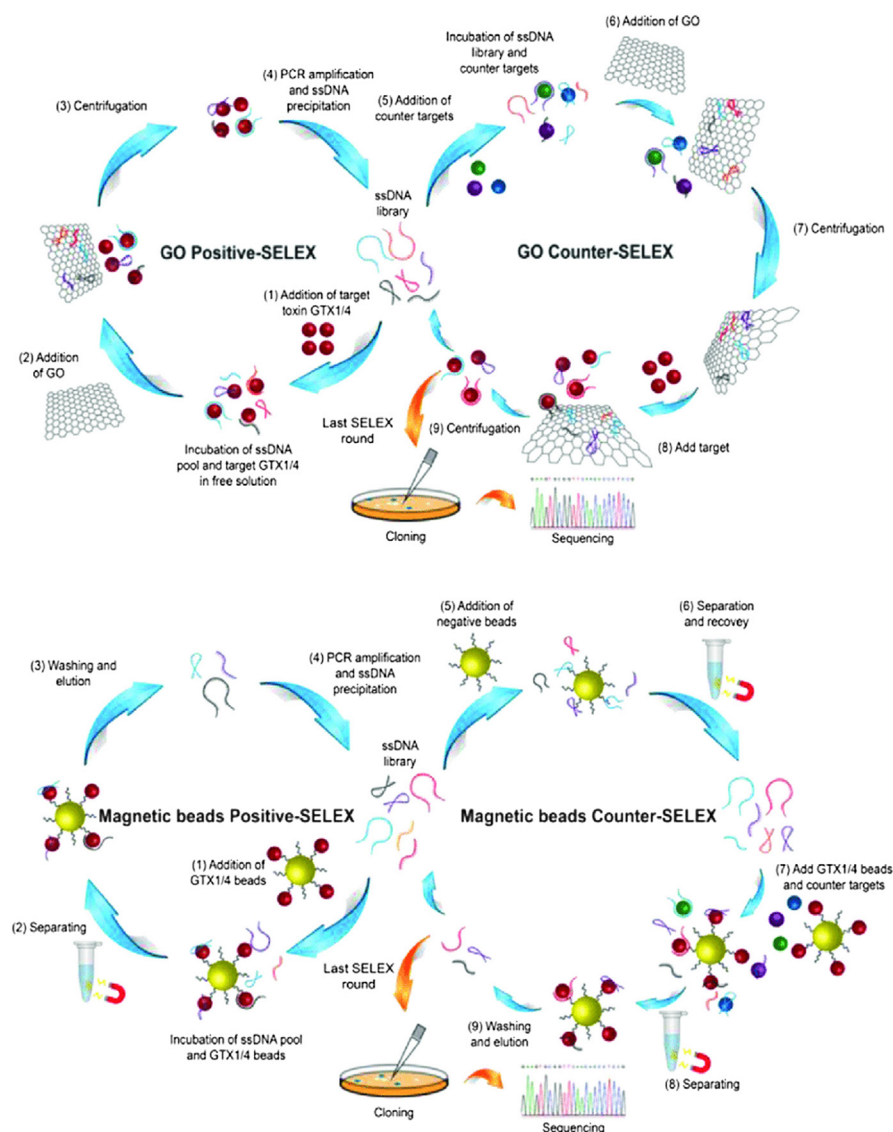
## 30.8 Aptamers

Aptamers are relatively a new class of single-stranded nucleic acid molecules that bind a ligand with high affinity and high specificity [19]. The term aptamer is derived from the Latin word *aptus* meaning “to fit” and the Greek word *meros* for “region.” Aptamers are created from a large selection of oligonucleotides with random sequences. Aptamer molecules have the potential of multiple therapeutic uses as these can bind to multiple target molecules including metal ions, organic compounds, nucleotides, amino acids, peptides, and proteins, including proteins on the surface of viruses or cells [20].

Most aptamers are around 15–50 nt long, however, these could be up to 100 nt in length as well. Depending on the length, an aptamer could be designed as ligands for large molecules, binding to equal size or small molecules for specific receptors. In contrast to other single-stranded nucleic acid molecules, like AS ONs, ribozymes and siRNAs, aptamers bind their target by Watson–Crick base pairing by folding into complex three-dimensional structures and recognize specific targets by steric interactions. In comparison to AS ONs, aptamers have the advantage of binding to extracellular targets, for example, coagulation factors or cytokines. Some of these are already evaluated for clinical applications.

Aptamers are usually assembled from a large oligonucleotides library by an in vitro selection procedure known as SELEX (Systemic Evolution of Ligands by EXponential enrichment) (Fig. 30.10). It is important for aptamers to be stabilized against nucleolytic degradation to ensure adequate biological activity. This is accomplished by introducing chemically modified nucleotides either during the selection procedure or incorporation of post-SELEX site-specific modification. A new class of nucleases resistant aptamers, called Spiegelmers, is now developed consisting of L-RNA enantiomer.

Two specific classes of aptamers are evaluated for clinical use. These include decoy oligonucleotides and immunostimulatory oligonucleotides. The decoy oligonucleotides are double stranded and can bind to transcription factors, preventing them from translocating into the nucleus and inducing transcription. The immunostimulatory oligonucleotides contain CpG (cytosine-phosphate-guanosine) motifs. They bind to toll-like receptor 9 and activate an immune response. These are being used as vaccine adjuvants and immunotherapy for allergy, cancer, and infectious diseases [20].



**FIGURE 30.10** SELEX—the in vitro method of selecting high affinity aptamers. SELEX, Systemic Evolution of Ligands by EXponential enrichment. Source: Adapted with permission from Gao, S, Hu, B, Zheng, X, Cao, Y, Liu, D, Sun, M, et al., Gonyautoxin 1/4 aptamers with high-affinity and high-specificity: From efficient selection to aptasensor application. *Biosensors and Bioelectronics*, 2016;79:938–44 [21].

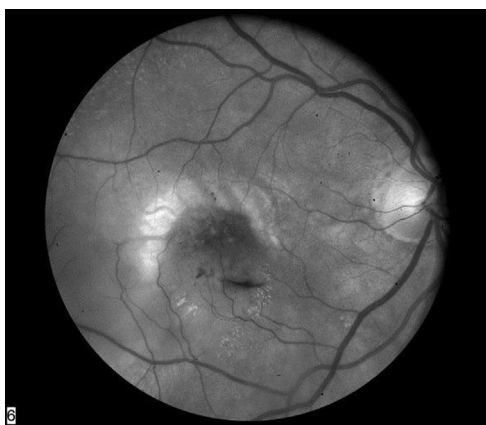
One of the major clinical applications of aptamers is the enhancement of vaccine efficacy. Among the most advanced is the HEPLISAV, a modified hepatitis B surface antigen that is aimed for individuals who respond poorly to conventional hepatitis B vaccination. Large clinical trials have demonstrated the immunostimulatory activity of adjuvant CpG oligonucleotides. This is also shown to maintain the vaccine efficacy and is particularly useful for use during pandemic when supplies of the vaccine could be very low [22].

There are a number of clinical applications of aptamers and are going through different phases of clinical trials. An important example is the wet age AMD. The mean visual acuity of patients treated with pegaptanib, approved by the FDA, is significantly improved compared to the standard care (Fig. 30.11) [23].

In summary the development of aptamers, decoy oligonucleotides, and immunostimulatory oligonucleotides shows that nucleic acids can be used not only to bind RNA or DNA molecules by standard Watson–Crick base pairing but can also function through targeting with extracellular peptides or proteins [24].

### 30.9 Gene and genome editing

Amongst many approaches of genome-based targeted molecular therapeutics are discussed in this chapter, gene/genome editing is by far the most challenging and promising. This method allows precise manipulation of



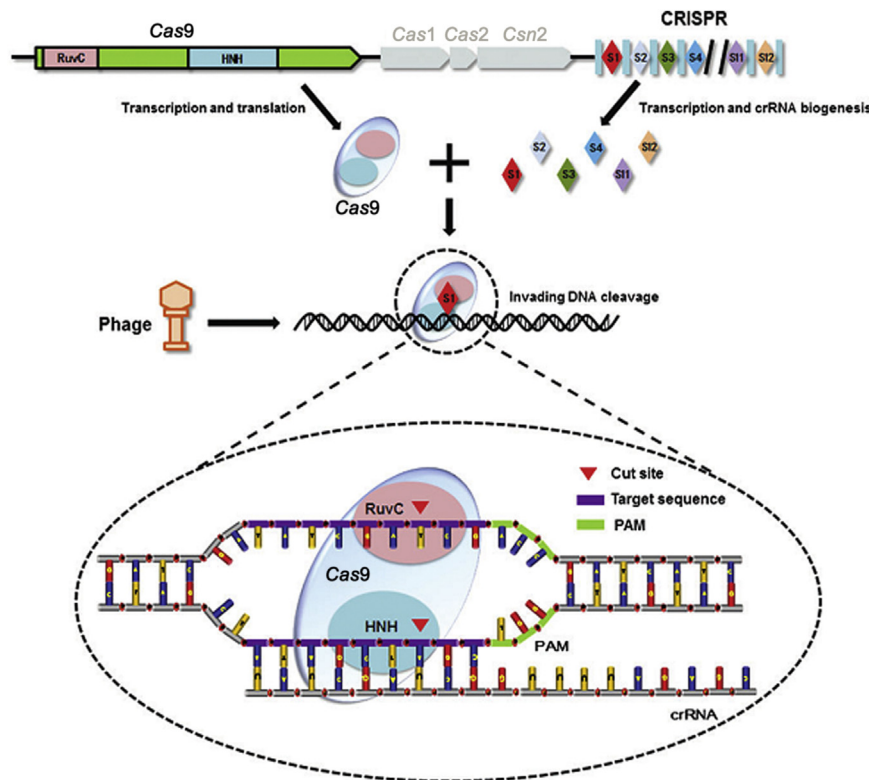
**FIGURE 30.11** Improvement of visual acuity following treatment with intraocular injection of pegaptanib in adult onset macular degeneration. Source: Adapted with permission from Singh, RP, Sears, JE, *Retinal Pigment Epithelial Tears After Pegaptanib Injection for Exudative Age-related Macular Degeneration*. *American Journal of Ophthalmology* 2006;142(1):160–2 [25].

specific gene or selected genomic elements for achieving the desired gene function, thus facilitating the treatment of a genetic disease or altering the biological function. In this context, techniques involving the “Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) protein 9 system” provide a robust, multiplex, and versatile genome editing tool [26].

CRISPR/Cas9 comprises a nonspecific Cas9 nuclease and a set of programmable sequence-specific CRISPR RNA (crRNA), which can guide Cas9 to cleave DNA and generate double-strand breaks (DSBs) at target sites. Subsequent cellular DNA repair process leads to desired insertions, deletions, or substitutions at target sites. The specificity of CRISPR/Cas9-mediated DNA cleavage requires target sequences matching crRNA and a protospacer adjacent motif (PAM) locating at downstream of target sequences. Multiple artificial nuclease systems have been developed for genome editing. Zinc-finger nucleases (ZFNs) are one of widely applied engineered nucleases. ZFNs contain a common Cys<sub>2</sub>-His<sub>2</sub> DNA-binding domain and a DNA cleavage domain of the FokI restriction endonuclease. Another popular genome editing platform is transcription activator-like effector nucleases (TALENs), which are derived from a natural protein of plant pathogenic bacteria *Xanthomonas*. The DNA-binding domain of TALENs is composed of 33–35 conserved amino acid repeated motifs, each of which recognizes a specific nucleotide. Through shuffling repeated amino acid recognition motifs, TALENs can be programmed to target-specific DNA sequence. Recently, CRISPR/CRISPR-associated (Cas) protein 9 system provides an alternative to ZFNs and TALENs for genome editing [27]. Distinct from the protein-guided DNA cleavage of ZFNs and TALENs, CRISPR/Cas9 depends on small RNA for sequence-specific cleavage. Because only programmable RNA is required to generate sequence specificity, CRISPR/Cas9 is easily applicable and develops very fast over the past year.

CRISPR/Cas9-mediated genome editing depends on the generation of DSB and subsequent cellular DNA repair process. In endogenous CRISPR/Cas9 system, mature crRNA is combined with transactivating crRNA (tracrRNA) to form a tracrRNA:crRNA complex that guides Cas9 to a target site. TracrRNA is partially complementary to crRNA and contributes to crRNA maturation. At the target site, CRISPR/Cas9-mediated sequence-specific cleavage requires a DNA sequence protospacer matching crRNA and a short PAM. After binding to the target site the DNA single-stranded matching crRNA and opposite strand are cleaved, respectively, by the HNH nuclease domain and RuvC-like nuclease domain of Cas9, generating a DSB at the target site [19]. A delicate guide RNA (gRNA), which was a chimeric RNA containing all essential crRNA and tracrRNA components is developed for enhancing the efficiency of genome editing. Multiple CRISPR/Cas9 variants have been developed, recognizing 20 or 24 nt sequences matching engineered gRNA and 2–4 nt PAM sequences at target sites. Therefore CRISPR/Cas9 can theoretically target a specific DNA sequence with 22–29 nt, which is unique in most genomes. However, recent studies observed that CRISPR/Cas9 had high tolerance to base pair mismatches between gRNA and its complementary target sequence, which was sensitive to the numbers, positions, and distribution of mismatches [28] (Fig. 30.12).

A number of biological applications of CRISPR/Cas9 genome editing system are possible. Current techniques allow targeting specific gene and as well-multiple genomic sites. This has been achieved in both mammalian and



**FIGURE 30.12** Schematic of CRISPR/Cas9-mediated DNA cleavage. Mature crRNA guides Cas9 to the target site of invading phage DNA. The DNA single-strand matching crRNA and opposite strand are cut, respectively, by the HNH nuclease domain and RuvC-like nuclease domain of Cas9, generating a DSB at the target site. The specificity of CRISPR/Cas9-mediated DNA cleavage requires target sequence matching crRNA and a 3 nt PAM locating at downstream of the target sequence. *crRNA*, Clustered regularly interspaced short palindromic repeats RNA; *DSBs*, double-strand breaks; *PAM*, protospacer adjacent motif. Source: Adapted with permission from *Hum Mol Genet* 2014;23(1):R42, Oxford University Press.

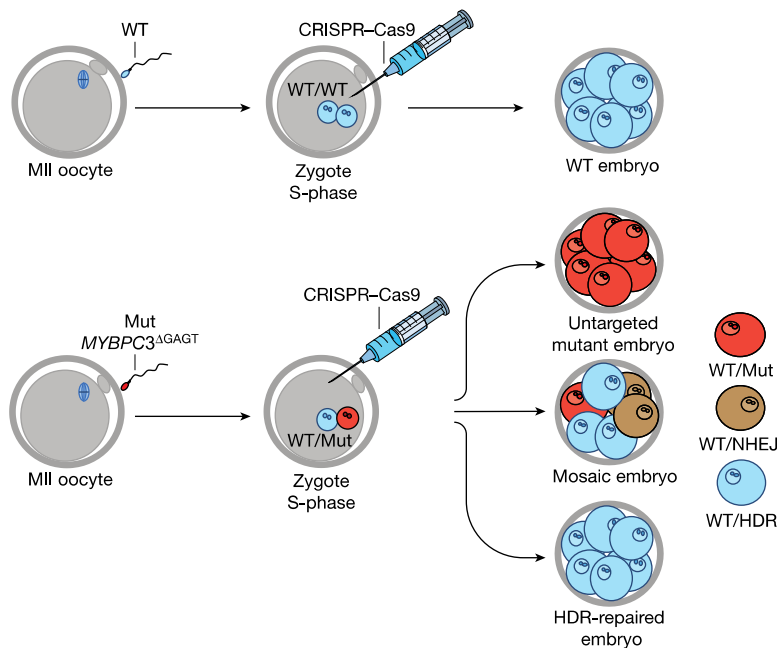
nonmammalian genomes. An important application of the CRISPR/Cas9 genome editing allows desired phenotype modification. In the case of a disease, biological resistance of threshold enhancement could be possible. This is evident from recent successes in plants and crop research. Examples include transfer of green fluorescence protein gene into *Arabidopsis* and tobacco genomes, bacterial blight susceptibility genes into rice genome, and modification of crop genomes to improve the crop quality [29].

There are many challenging clinical applications of the CRISPR/Cas9 genome editing system. The precision genome editing has the potential to permanently cure diseases through disrupting endogenous disease-causing genes, correcting disease-causing mutations or inserting new protective genes [30,31]. In the past, limited successes of genome editing were possible using ZFNs for correcting the gene mutations causing sickle-cell disease, hemophilia B, inducing resistance to virus infection in human cells, and enhancing the efficiency of immunotherapies. The CRISPR/Cas9 technique also has the potential of correcting pathogenic mutation or polymorphic variant in a complex disease, like Parkinson's disease [31].

With the rapid development of induced pluripotent stem (iPS) cells technology, engineered nucleases are applied to genome manipulation of iPS cells. The unlimited self-renewing and multipotential differentiation capacity of iPS cells make them very useful in disease modeling and gene therapy. Using CRISPR/Cas9, an iPS cell model for immunodeficiency, centromeric region instability, and facial anomalies syndrome (ICF) caused by DNMT3B gene mutation was created [32]. However, it is still far too early to design any clinical application.

Recently, Ma et al. reported the applications of CRISPR/Cas9 gene editing method for the correction of the heterozygous pathogenic *MYBPC3* in a family with autosomal dominant hypertrophic cardiomyopathy (HCM) [33]. The index patient was an adult male patient with well-documented familial HCM caused by a heterozygous dominant 4-bp GAGT deletion (g.9836\_9839 del., NC\_000011.10) in exon 16 of *MYBPC3*. He was managed with an implantable cardioverter defibrillator and prescribed a combination of beta blockers and antiarrhythmic medications. Skin fibroblast cultures were expanded and used to generate heterozygous patient iPSCs. Two single-gRNA–Cas9 constructs were designed to target this specific *MYBPC3*ΔGAGT deletion (Fig. 30.13) along with two exogenous single-stranded oligodeoxynucleotide (ssODN) templates encoding homology arms to the targeted region. To differentiate from the wild-type allele, two synonymous single-nucleotide substitutions were introduced into each ssODN template. In addition, ssODN-2 nucleotide substitutions provided an additional restriction enzyme (BstBI) recognition site.





**FIGURE 30.13** MYBPC3 gene correction in S-phase human embryos using CRISPR/Cas9 method; schematic of MYBPC3<sup>ΔGAGT</sup> gene targeting by injection of CRISPR–Cas9 into human zygotes at the S-phase of the cell cycle. MII oocytes were fertilized by sperm from a heterozygous patient with equal numbers of mutant and WT spermatozoa. CRISPR–Cas9 was then injected into one-cell zygotes. Embryos at the 4–8-cell stage were collected for genetic analysis. Injection during S-phase resulted in mosaic embryos consisting of nontargeted mutant, targeted NHEJ-repaired, and targeted HDR-repaired blastomeres. WT, Wild-type. Source: Adapted with permission from Ma H, et al. Correction of a pathogenic gene mutation in human embryos. *Nature* 2017;548:413–9; doi:10.1038/nature23305, Figure 1.

Following the successful application of CRISPR/Cas9 gene editing in patient's iPSCs, this method was applied in the preimplantation zygote. It was achieved with precise CRISPR–Cas9-based targeting accuracy and high homology-directed repair efficiency by activating an endogenous, germline-specific DNA repair response. Induced DSBs at the mutant paternal allele were predominantly repaired using the homologous wild-type maternal gene instead of a synthetic DNA template. By modulating the cell cycle stage at which the DSB was induced, researchers were able to avoid mosaicism in cleaving embryos and achieve a high yield of homozygous embryos carrying the wild-type MYBPC3 gene without evidence of off-target mutations. The efficiency, accuracy, and safety of the approach presented have provided evidence that it has potential to be used for the correction of heritable mutations in human embryos by complementing preimplantation genetic diagnosis. However, much remains to be considered before clinical applications, including the reproducibility of the technique with other heterozygous mutations.

### 30.10 Summary

This chapter provides the progress made so far and the landscape of *gene and genome-based* molecular therapeutics. Earlier progress in the recombinant DNA technology remains a major advance. Newer methods in the stem cell technology revolutionized the prospect and potential of stem-cell regeneration used in treating a number of clinical conditions. To large extent the stem cell–based cellular regeneration facilitated treatment in many disorders where attempts using the gene therapy had failed. Nevertheless, gene therapy remains a unique paradigm that remains to be fully exploited. The progress made in the last decade led to the much deeper understanding of the therapeutic applications of oligonucleotides, RNAi, and aptamers. These are continually exploited in finding novel therapeutic models undergoing clinical trials. The highlight of advances in this field is undoubtedly the CRISPR/Cas9-based gene and genome editing system. Earlier indications are not limited to therapy of human disease but have the potential widespread use in many other biotechnological fields including animal breeding, horticulture, and agriculture. Recent reports on crop improvement using the gene/genome editing are promising.

### Disclaimer and acknowledgments

The author has compiled this chapter using information from a number of published resources accessible on the web. The request for permissions has been made to respective resources for including figures with clearly



identified citation. The author wishes to acknowledge authors, contributors, and publishers of these resources for permitting use in this chapter solely for educational purpose. The author does not, in any manner, impose or reflect ownership of the material used.

## References

- [1] Walsh G. Biopharmaceutical benchmarks 2014. *Nat Biotechnol* 2014;32:992–1000.
- [2] Veronese FM, Pasut G. PEGylation: Posttranslational bioengineering of protein biotherapeutics. *Drug Discovery Today: Technologies* 2008;5(2–3):e57–64.
- [3] MacPherson G, Austyn J. Exploring immunology: concepts and evidence. Weinheim: Wiley-VCH Verlag GmbH; 2012.
- [4] Robinson DA, Daley GQ. The promise of induced pluripotent stem cells in research and therapy. *Nature* 2012;481:295–305.
- [5] Yamanaka S. Strategies and New Developments in the Generation of Patient-Specific Pluripotent Stem Cells. *Cell Stem Cell* 2007;1(1):39–49.
- [6] Gattegno-Ho D, Argyle S, Argyle DJ. Stem cells and veterinary medicine: Tools to understand diseases and enable tissue regeneration and drug discovery. *The Veterinary Journal* 2012;191(1):19–27.
- [7] Quiskamp N, Bruin JE, Kieffer TJ. Differentiation of human pluripotent stem cells into  $\beta$ -cells: Potential and challenges. *Best Practice & Research Clinical Endocrinology & Metabolism* 2015;29(6):833–47.
- [8] Ibraheem D, Elaissari A, Fessi H. Gene therapy and DNA delivery systems. *International Journal of Pharmaceutics* 2014;459(1–2):70–83.
- [9] Wong GK, Chiu AT. Gene therapy, gene targeting and induced pluripotent stem cells: applications in monogenic disease treatment. *Biotechnology Advances* 2011;29(1):1–10.
- [10] Aartsma-Rus A, Janson AA, Kaman WE, Bremmer-Bout M, van Ommen GJ, den Dunnen JT, et al. Antisense-Induced Multiexon Skipping for Duchenne Muscular Dystrophy Makes More Sense. *Am J Hum Genet* 2004;74(1):83–92.
- [11] Kurreck J. Antisense technologies improvement through novel chemical modifications. *Eur J Biochem* 2003;270:1628–44.
- [12] Cech TR. Structural biology: the ribosome is a ribozyme. *Science* 2000;289:878–9.
- [13] Schubert S, Kurreck J. Ribozyme- and deoxyribozyme-strategies for medical applications. *Curr Drug Targets* 2004;. Available from: <https://doi.org/10.2174/1389450043345092>.
- [14] Vaucheret H, Béclin C, Fagard M. Post-transcriptional gene silencing in plants. *J Cell Sci* 2001;114(Pt 17):3083–91.
- [15] Kurreck J. RNA interference: from basic research to therapeutic applications. *Angew Chem Int Ed* 2009;48:1378–98.
- [16] Jeong E, Kim H, Jang B, Cho H, Ryu J, et al. Technological development of structural DNA/RNA-based RNAi systems and their applications. *Advanced Drug Delivery Reviews* 2016;104:29–43.
- [17] Kleinman M, et al. Short acting RNAi induce retinal degeneration via TLR3 and IRF3. *Mol Ther* 2012;20(1):101–8.
- [18] Urakami T, Oku N. Current status of siRNA delivery technology and siRNA drug development. *Open Drug Deliv J* 2007;2007(1):20–7.
- [19] Famulok M, Mayer G. Aptamer molecules as sensors and detectors. *Acc Chem Res* 2011;44:1349–58.
- [20] Keefe AD, Pai S, Ellington A. Aptamers in therapeutics. *Nat Rev Drug Discov* 2010;5:123–32.
- [21] Gao S, Hu B, Zheng X, Cao Y, Liu D, Sun M, et al. Gonyautoxin 1/4 aptamers with high-affinity and high-specificity: From efficient selection to aptasensor application. *Biosensors and Bioelectronics* 2016;79:938–44.
- [22] Zhou J, Bobbin ML, Burnett RR, et al. Current progress of RNA aptamer-based therapeutics. *Front Genet* 2012;3:234.
- [23] Ng EW, Shima DT, Calias P, et al. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov* 2006;5:123–32.
- [24] Koppikar P, Bromberg J. State-of-the-art approach using oligonucleotide decoys to target ‘undruggable’. *Cancer Discov* 2012;2:670–2.
- [25] Singh RP, Sears JE. Retinal Pigment Epithelial Tears After Pegaptanib Injection for Exudative Age-related Macular Degeneration. *American Journal of Ophthalmology* 2006;142(1):160–2.
- [26] Zhang F, Wen Y, Guo X. CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum Mol Genet* 2014;23(1):R40–6. Available from: <https://doi.org/10.1093/hmg/ddu125>.
- [27] Pennisi E. The CRISPR craze. *Science* 2013;341:833–6 19.
- [28] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339:819–23.
- [29] Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res* 2013;41:e188.
- [30] Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood* 2011;118:4599–608.
- [31] Soldner F, Laganier J, Cheng AW, Hockemeyer D, Gao Q, Alagappan R, et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 2011;146:318–31.
- [32] Horii T, Tamura D, Morita S, Kimura M, Hatada I. Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. *Int J Mol Sci* 2013;14:19774–81.
- [33] Ma H, et al. Correction of a pathogenic gene mutation in human embryos. *Nature* 2017;548:413–19. Available from: <https://doi.org/10.1038/nature23305>.

# Personalizing medicine with pharmacogenetics and pharmacogenomics

N.J. Gogtay, U.M. Thatte, R. Ravi and D. Bose

Department of Clinical Pharmacology, Seth GS Medical College and KEM Hospital, Mumbai, India

*Give different drugs to different patients, for the sweet ones do not benefit everyone, nor do the astringent ones, nor are all the patients able to drink the same things.* Hippocrates, circa 400 BC.

*If it were not for the great variability among individuals, medicine might as well be a science, not an art.* Sir William Osler, 1892.

*Knowledge of sequences could contribute much to our understanding of living matter.* Frederick Sanger, circa 1977.

## 31.1 Introduction and historical perspective

Treatment with drugs or pharmacotherapy remains the cornerstone approach to the management of disease. The use of drugs is, however, fraught with several challenges not the least of which is wide interindividual variation in response. Evidence puts efficacy at 30%–60% for several drugs, and 7% of patients are known to experience serious adverse reactions [1]. The process of variation in drug response is usually addressed by the treating physician using a trial and error process.

It was the Greek philosopher and mathematician Pythagorus in 510 BC who recorded the first interindividual difference when he noted that some patients developed hemolytic anemia after ingestion of the fava bean, while others did not [2]. The term “pharmacogenetics” was coined by the German physician Friedrich Vogel in 1957 although a clear-cut definition was provided only in 1962 by Kalow as “study of heredity and response to drugs” [3]. Pharmacogenetics can be viewed as a scientific discipline that arose from the confluence of genetics, biochemistry, and pharmacology [4]. Table 31.1 chronologically lists the seminal discoveries in this area [5].

A field that came after pharmacogenetics was that of pharmacogenomics. The latter field resulted from several advances in molecular medicine. Although the two terms are frequently used interchangeably, *pharmacogenetics* can be understood as the study of drug response in relation to a specific gene or a set of genes, whereas *pharmacogenomics* is much broader and can be defined as the science of the study of drug response in relation to the entire genome [6].

The United States Food and Drug Administration (US-FDA) defines “pharmacogenetics” as the study of variations in DNA *sequence* as related to drug response, while pharmacogenomics is defined as “the study of variations of DNA and RNA *characteristics* as related to drug response” [7]. Pharmacogenomics thus is all encompassing and addresses the impact of genomic information on the drug-discovery process, identification of candidate genes and polymorphisms, correlation of polymorphisms with therapies, prediction of drug response (both benefit and harm), clinical outcomes, minimizing adverse events, and the use of genotype-guided dosing regimens. This is because pharmacogenomics looks at molecular determinants at the genome-, transcriptome-, and proteome-wide levels, whereas pharmacogenetics is limited to one or more specific genes or genetic markers.

Clinical DNA-based testing for disease detection began for the diagnosis of sickle-cell disease (Hb S mutation in the  $\beta$ -globin gene; the first inherited disease identified at the molecular level) [8] and was initially a niche specialty that catered to rare disorders. Subsequently, the genetic basis of many disorders was identified. Today, the testing spans several medical disciplines and includes newborn screening, testing for disorders of inheritance,

**TABLE 31.1** Historical overview of pharmacogenetics and pharmacogenomics.

Year	Individual(s)	Landmark/event
510 BC	Pythagoras	Recognition of the danger of ingesting fava beans (G6PD deficiency)
1866	Mendel	Establishment of the rules of heredity
1906	Garrod	Publication of Inborn Errors of Metabolism
1932	Snyder	Characterization of “phenylthiourea nontaster” as an autosomal recessive trait
1956	Carson et al.	Discovery of G6PD deficiency
1957	Motulsky	Further refined the concept that inherited defects of metabolism may explain individual differences in drug response
1957	Kalow and Genest	Characterization of serum cholinesterase deficiency
1957	Vogel	Coined the term pharmacogenetics
1960	Price and Evans	Characterization of acetylator polymorphism
1962	Kalow	Publication of “Pharmacogenetics—Heredity and the Response to Drugs”
1977/79	Mahgoub et al. and Eichelbaum et al.	Discovery of the polymorphism in debrisoquine hydroxylase sparteine oxidase
1988	Gonzalez et al.	Characterization of genetic defect in debrisoquine hydroxylase, later termed CYP2D6
1988–2000	Various scientists	Identification of specific polymorphisms in various phase I and phase II drug-metabolizing enzymes and later in drug transporters
2000	Public–private partnership	Completion of the first draft of the human genome
2000	The International SNP Map Working Group	Completion of the map of human genome sequence variation containing 1.42 million single-nucleotide polymorphisms

*Reproduced with permission from the Br J Clin Pharmacol.*

predictive testing including pharmacogenetic testing and includes population-based screening programs in several countries (breast and prostate cancer screening for example).

Both pharmacogenetics and pharmacogenomics are important areas for both scientific research and diagnosis and therapy as the promise of targeting drugs according to an individual’s specific genetic makeup of each patient offers great promise and can produce tangible benefits in clinical practice and health policy.

## 31.2 The Human Genome Project

The discovery of the double-helical structure of DNA in 1953 by Watson and Crick laid the foundation stone for what we know today as genomic medicine [9]. In the mid-70s, Frederick Sanger developed techniques to sequence DNA [10]. This was followed in the 1980s by developments in automation and gave birth to the idea that the entire human genome could perhaps be sequenced. The Human Genome Project (HGP) was an international collaborative research project that aimed to identify the sequence of the human genome and the genes contained therein. This project was coordinated by the National Institutes of Health (NIH) and the US Department of Energy. In addition, contributions were made from universities across the United States. Other partners included researchers from the United Kingdom, France, Germany, Japan, and China. The project ran from 1990 to 2003 and was completed 2 years ahead of schedule [11]. In June 2000 came the announcement that the majority of the human genome had in fact been sequenced, which was followed by the publication of 90% of the sequence of the genome’s 3 billion base pairs in the journal *Nature*, in February 2001. The advent and deployment of improved research techniques, including the use of restriction fragment-length polymorphisms, the polymerase chain reaction, bacterial and yeast artificial chromosomes, and pulsed-field gel electrophoresis enabled rapid progress of the project [12].

The entire genome has 3 billion base pairs and about 30,000 genes; a lot fewer than originally envisaged. The HGP has helped physicians and researchers to build upon the knowledge, resources, and technology

that has emanated from the project to understand the contributions of genetics to health and disease. This gave birth to the field of genomic medicine and now the widespread use of genetics/genomics in diagnosis, monitoring, and treatment of disease. At the completion of the HGP, Dr. Francis Collins, the director of the project emphasized several prominent expectations from the project. These included targeted drug discovery, using genetic information to predict response to therapy and also projected that the pharmacogenomics approach for predicting response to treatment would become standard of care by the year 2020 [13].

### 31.3 The introduction of the field of personalized medicine/precision medicine

The mapping of the human genome led to the understanding that mankind's genetic makeup is 99% identical, and it is the 1% that is responsible for the differences. The mapping also offered hope to "individualize" or "personalize" treatment, a term we call today as "personalized medicine" (PM). The concept of "personalized medicine" was anticipated in the late 1800s by the Canadian physician Sir William Osler who noted "the great variability amongst individuals." The term itself first appeared in literature in the late 1990s and is used widely today though it is interpreted to mean different things by different people. A commonly used definition is "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person." [14]. The National Cancer Institute (of the National Institute of Health, the United States) has a based and an all-encompassing definition that states "personalized medicine is a form of medicine that uses information about a person's genes, proteins and environment to prevent, diagnose and treat disease." It is also called precision medicine or individualized medicine [15]. Thus pharmacogenetics and pharmacogenomics can be viewed as a part of the field of precision medicine. They can also be viewed as one of the most "actionable" areas of the precision medicine paradigm. The idea is to customize health care, with decisions and treatments tailored to each individual patient in every way possible. While the term itself is relatively new, the concept is in use already. A classic example of PM is that of tailoring the blood group to the individual patient prior to a blood transfusion to minimize or prevent the risk of transfusion reactions.

The progress in the field of PM can be attributed to one key discovery and one key invention—(1) the identification of single-nucleotide polymorphisms (SNPs) genotyping [11] and (2) microarray/biochips [16]. SNPs are essentially single-nucleotide changes in the entire genome that occur in more than 1% of the population. Their delineation is important as they determine both disease susceptibility and response to therapy. The invention of the microarray with its ability to rapidly analyze a patient's genome in its entirety as also to store it revolutionized the field of PM. Today, the NIH along with other research centers has begun a Precision Medicine Initiative that has both short-term and long-term goals. The former involves the application of precision medicine in the area of cancer also called cancer precision medicine. The latter essentially involves the application of precision medicine to all areas of health care on a large scale.

An example of the practice of precision medicine in oncology can be seen in the area of lung cancer. It is now known that lung cancer is a heterogeneous disease with distinct subtypes that differ in terms of morphology, molecular characteristics, and therapeutic vulnerability [17]. Up until recently, lung cancer was divided into only two broad categories—non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC), and these could be further subdivided histologically into adenocarcinoma (~45%), squamous-cell carcinoma (~25%), large-cell carcinoma (10%), and small-cell carcinoma (SCLC, ~20%). Two molecular alterations that now drive therapy in lung cancer include the epidermal growth factor receptor (EGFR) mutations and anaplastic lymphoma kinase (ALK) gene rearrangement [18]. The presence of these two molecular alterations confers sensitivity to targeted therapies such as the tyrosine kinase inhibitors such as erlotinib (for patients with EGFR mutations) and ALK inhibitors such as crizotinib for patients with ALK rearrangements [18]. Treatment with specific inhibitors in these individuals is superior to chemotherapy. The converse is also true. Patients that lack the EGFR mutations or ALK rearrangements are better candidates for chemotherapy.

### 31.4 Genetic and molecular basis of the individual drug-response variation

Human variation can be judged by the way they respond to drugs, in terms of both effective response and minimal side effects. To large extent, while this variation is probably due to clinical and environmental factors,

individual genomic variation is of basic importance. In this context, two fundamental concepts are important—pharmacokinetics and pharmacodynamics. Pharmacokinetics is the study of “what the body does to a drug.” This can be divided into four different processes: absorption, distribution, metabolism, and excretion. A number of genes and metabolizing enzymes influence pharmacokinetics.

Pharmacodynamics is the study of “what the drug does to the body.” Pharmacodynamic sources of variation in drug action can be due to changes in the expression of the drug target, their affinity or their selectivity. This can be caused by genetic factors and also by disease. For example, in renal impairment, pharmacodynamic factors determine the sensitivity to a drug.

Age, body weight, renal and hepatic function, concomitant drugs, comorbidities, diet, smoking, and alcohol are all known to affect the pharmacokinetics and sometimes the pharmacodynamics of drugs. Genetic factors affecting both the pharmacokinetics and pharmacodynamics of drugs also play a role in determining how individuals respond to drugs.

### 31.4.1 Genetic factors in pharmacokinetics

Interindividual variation is important for both presystemic hepatic metabolism and the rate of systemic hepatic clearance. Variability in the genes that encode drug-metabolizing enzymes (Table 31.2) is a major determinant of the interindividual differences in the therapeutic and adverse responses to drug treatment. Genetic polymorphisms can affect both phase I and phase II drug-metabolizing enzymes. *Phase I* is the modification of a drug, by oxidation, reduction, or hydrolysis; oxidation is the most frequent route and is largely undertaken by a family of isoenzymes known as the cytochrome P450 system. Inhibition or induction of cytochrome P450 isoenzymes is a major cause of drug interactions. *Phase II* involves conjugation with glucuronate, sulfate, acetate, or other substances to render the drug more water-soluble and therefore able to be excreted in the urine.

Approximately 25% of all medicines currently in use are substrates for CYP2D6. The frequencies of the variant alleles show racial variation, and a small proportion of individuals may have three or more copies of the active gene. The phenotypic consequences of the defective CYP2D6 include the increased risk of toxicity with those antidepressants or antipsychotics that undergo metabolism by this pathway. In addition, reduced conversion of a prodrug to the active metabolite may also compromise efficacy in some patients; for example, poor metabolizers of CYP2D6 may be at increased risk of breast cancer relapse with tamoxifen. Conversely, in individuals with multiple copies of the active gene (ultrarapid metabolizers), there are extremely rapid rates of metabolism and therapeutic failure at conventional doses. Ultrarapid metabolizers may also be at increased risk of toxicity with the conversion of prodrugs to active metabolites; this has been shown with codeine, which is metabolized to morphine and has caused respiratory depression, particularly in children.

Warfarin is predominantly metabolized by CYP2C9. In most populations, between 2% and 10% are homozygous for an allele that results in low enzyme activity. Such individuals will therefore metabolize warfarin more slowly, leading to higher plasma levels, a greater risk of bleeding, and a requirement for lower doses if the international normalized ratio (INR) is to be maintained within the therapeutic range. When combined with the genetic polymorphisms in the vitamin K epoxide reductase complex (*VKORC1*) genes (a pharmacodynamic variation), and age and body mass index, over 50% of the variation in individual daily dose requirement can be predicted. Preprescription genotyping for CYP2C9 and *VKORC1* has been shown to improve anticoagulation control.

Individual differences in the activity of thiopurine methyltransferase, a phase II enzyme, are used to determine the appropriate doses of mercaptopurine and azathioprine. Testing for TMPT activity is therefore undertaken routinely in children undergoing treatment for acute lymphatic leukemia and people with Crohn's disease. Slow acetylators (deficient in *N*-acetyltransferase type 2) have an increased risk of hepatotoxicity with isoniazid, while individuals with deficient *UGT1A1* activity are at increased risk of toxicity from irinotecan.

### 31.4.2 Genetic factors in pharmacodynamics

Genetic variation also affects pharmacodynamic targets, which leads to variation in both the efficacy and the safety of drugs. Key examples include variation in the  $\beta_2$ -adrenoceptor gene, which can affect response to salbutamol (albuterol). Glucose-6-phosphate dehydrogenase deficiency, the most common enzyme deficiency in the world, can predispose patients to acute red cell hemolysis with certain drugs such as primaquine, dapsone, sulphonamides, and rasburicase. The biggest clinical advances in this area have occurred in the following two therapeutic areas:



**TABLE 31.2** Genes and genetic polymorphisms in drug response.

Organ or system involved	Associated gene/allele	Drug/drug response phenotype	Pharmacokinetic/Pharmacodynamic
<b>Blood</b>			
Red blood cells	<i>G6PD</i>	Primaquine and others	Pharmacodynamic
	<i>TPMT*2</i>	Azathioprine/6MP-induced neutropenia	Pharmacokinetic
Neutrophils	<i>UGT1A1*28</i>	Irinotecan-induced neutropenia	Pharmacokinetic
Platelets	<i>CYP2C19*2</i>	Stent thrombosis	Pharmacokinetic
Coagulation	<i>CYP2C9*2, *3, VKORC1</i>	Warfarin dose requirement	Mixed
<b>Brain and peripheral nervous system</b>			
CNS depression	<i>CYP2D6*N</i>	Codeine-related sedation and respiratory depression	Pharmacokinetic
Anesthesia	Butyrylcholinesterase	Prolonged apnea	Pharmacokinetic
Peripheral nerves	<i>NAT-2</i>	Isoniazid-induced peripheral neuropathy	Pharmacokinetic
<b>Drug hypersensitivity and liver injury</b>			
	<i>HLA-B*57:01</i>	Abacavir hypersensitivity	Pharmacodynamic
	<i>HLA-B*15:02</i>	Carbamazepine-induced Stevens – Johnson syndrome (in some Asian groups)	Pharmacodynamic
	<i>HLA-A*31:01</i>	Carbamazepine-induced hypersensitivity in Caucasians and Japanese	Pharmacodynamic
	<i>HLA-B*58:01</i>	Allopurinol-induced serious cutaneous reactions	Pharmacodynamic
	<i>HLA-B*57:01</i>	Flucloxacillin hepatotoxicity	Pharmacodynamic
<b>Malignancy</b>			
Breast cancer	<i>CYP2D6</i>	Response to tamoxifen	Pharmacokinetic
Chronic myeloid leukemia	<i>BCR–ABL</i>	Imatinib and other tyrosine kinase inhibitors	Pharmacokinetic
Colon cancer	<i>KRAS</i>	Cetuximab efficacy	Pharmacokinetic
Gastrointestinal stromal tumors	<i>c-kit</i>	Imatinib efficacy	Pharmacokinetic
Lung cancer	<i>EGFR</i>	Gefitinib efficacy	Pharmacokinetic
	<i>EML4-ALK</i>	Crizotinib efficacy	Pharmacokinetic
Malignant melanoma	<i>BRAF V600E</i>	Vemurafenib efficacy	Pharmacodynamic
<b>Muscle</b>			
General anesthetics	Ryanodine receptor	Malignant hyperthermia	Pharmacodynamic
Statins	<i>SLCO1B1</i>	Myopathy/Rhabdomyolysis	Pharmacokinetic

Adapted from Pirmohamed M. Pharmacogenetics: past, present and future. *Drug Discov Today* 2011;16:852–61.

**Cancer.** Sequencing of the cancer genome has identified novel driver mutations. This has led to the development of targeted therapies, which have been remarkably successful in the treatment of some malignancies, even when they are at an advanced stage; prominent examples include SCLC, colorectal cancer, and malignant melanoma.

**Drug safety.** Some immune-mediated reactions involving the skin and liver can now be predicted by genotyping for certain human leukocyte antigen (HLA) polymorphisms. For example, preprescription genotyping for *HLA-B\*57:01* has been shown to be clinically effective and cost-effective in preventing hypersensitivity to the anti-human immunodeficiency virus (HIV) drug abacavir. Similarly, *HLA-B\*15:02* predisposed patients from Southeast Asia to Stevens–Johnson syndrome (SJS) with carbamazepine (CBZ).

## 31.5 Tailoring or individualizing drug therapy—select examples of application of pharmacogenomics in clinical practice

### 31.5.1 *Trastuzumab and ERBB2 (HER2) genotype—tailoring treatment based on genomic testing*

The human EGFR (HER) family consists of four members: the EGFR, HER2, HER3, and HER4. All four are transmembrane tyrosine kinase receptors that regulate a number of important cellular processes including cell growth, survival, and differentiation. HER2 is a proto-oncogene. Proto-oncogenes are a group of genes that, when mutated or expressed at abnormally high levels, can contribute to normal cells becoming cancerous cells. The HER2 gene is overexpressed in a quarter of breast cancers and is also overexpressed in some cases of gastric cancer [19]. Tumors that express HER2 show accelerated growth and have a poorer prognosis. Testing for HER2 is done on the breast cancer tissue either by using immunohistochemistry or testing for gene amplification using in situ hybridization. HER2-targeted therapy with trastuzumab is recommended if HER2 test result is positive.

Trastuzumab received US-FDA approval in September 1998 and is typically used in these two cancers in the adjuvant setting though it can also be used in the neoadjuvant setting. Four major trials evaluated trastuzumab given as adjuvant therapy—Herceptin Adjuvant, National Surgical Adjuvant Breast and Bowel Project B-31, North Central Cancer Treatment Group N9831, and Breast Cancer International Research Group 006—including between them more than 13,000 women with HER-2-positive breast cancer. Between them, these trials investigated different adjuvant treatment approaches with trastuzumab [20–23]. All four trials showed a reduction in the risk of recurrence by half. The benefit was similar across the studies despite diversity in patient populations, chemotherapy regimens used, and the sequence of use of treatments. At present, trastuzumab is the standard of care for breast and gastric cancers that overexpress this protein and is often called the “poster child” of pharmacogenomics.

### 31.5.2 *Warfarin use as an anticoagulant—tailoring an individual’s dose using preprescription genetic information—testing for CYP2C9 and VKORC1*

Warfarin is the most widely prescribed oral anticoagulant worldwide. However, there is a greater than 10-fold interindividual variability in the dose required to attain a therapeutic response. The capricious nature of its response, the inherent complexity of the coagulation cascade, and the diversity of factors that affect its response (age, concomitant medications, vitamin K levels, and liver function) make effective anticoagulation extremely difficult. Warfarin is a racemic mixture of S and R warfarin, and CYP2C9 is the enzyme that primarily degrades S warfarin. The minor metabolic pathways include CYP2C8, CYP2C18, and CYP2C19. The S enantiomer is more potent and is largely responsible for the variation in the dosing requirements of the drug. Warfarin produces its effect by interfering with the synthesis of vitamin K–dependent clotting factors by inhibiting vitamin K epoxide reductase complex (VKORC1). Depletion of vitamin K leads to the production of nonfunctional clotting factors leading to anticoagulation.

Several variant alleles of CYP2C9 have been identified, and these lead both normal and reduced activity of the enzyme with the latter variant resulting in lower clearance of warfarin. Patients who have at least one copy of CYP2C9\*2 and CYP2C9\*3 have reduced metabolism and thus require lower daily doses of the drug relative to those who are homozygous for the wild-type CYP2C9\*1 allele [24]. With the VKORC1 gene, patients who have the –1639G > A polymorphism in the promoter region display greater sensitivity to warfarin and require lower doses [25]. One of the major benefits of using warfarin (among the many indications) is the reduction in risk of stroke in patients with non–rheumatic atrial fibrillation. The benefits of using warfarin are however offset by the potentially serious risks associated with this use that include both minor and major bleeding (including intracranial hemorrhage and death related to bleeding). The risk of bleeding is highest in the first few months of therapy and increases dramatically when the INR crosses 4.0. A retrospective audit of  $n = 555$  patients who presented to our warfarin clinic showed that the presence of a variant CYP2C9 allele (such as \*2 or \*3) was associated with twice the odds of bleeding (relative to those who did not have the allele); the presence of VKORC1 AG/AA haplotype was associated with three times the odds of bleeding (relative to those who did not have the haplotype), and the presence of both variants was associated with four times the odds of bleeding [26].

The FDA-approved warfarin drug label provides a dosing table based on CYP2C9 and VKORC1 genotypes [27] (Table 31.3). The label states if the patient’s CYP2C9 and/or VKORC1 genotype are known, the ranges listed in the table may be used to choose the initial dose. The label also states that patients with CYP2C9 \*1/\*3, \*2/\*2, \*2/

**TABLE 31.3** Individualizing Warfarin dosing based on patient's genetic information [27].

VKORC1	CYP2C9					
	*1/*1	*1/*2	*1/*3	*2/*2	*2/*3	*3/*3
<i>Dose in mg</i>						
GG	5–7	5–7	3–4	3–4	3–4	0.5–2
AG	5–7	3–4	3–4	0.5–2	0.5–2	0.5–2
AA	3–4	3–4	0.5–2	0.5–2	0.5–2	0.5–2

**TABLE 31.4** Human leukocyte antigen (HLA)-B\*57:01 carrier frequency range in diverse populations.

Population	HLA-B*57:01 carrier frequency range expressed as percentage
European	1.4–10.2
South American	1.1–3.1
African	0–3.2
Middle Eastern	0.5–6
Mexican	0–4.0
Asian	0–6.7
Southwest Asian (Indian)	3.8–19.6

\*3, and \*3/\*3 may require a longer time to achieve effective anticoagulation (as measured by time taken to achieve target INR).

### 31.5.3 Human leukocyte antigen testing in clinical practice

#### 31.5.3.1 Abacavir use in human immunodeficiency virus (HIV) infection—preventing an adverse effect through prescription human leukocyte antigen (HLA) testing

Abacavir is a nucleoside reverse transcriptase inhibitor used for the management of HIV infection. While, the drug is largely safe and effective, approximately 5% of patients receiving the drug are at risk of developing potentially life-threatening hypersensitivity reactions. It was discovered after several years of use of the drug that this reaction was immune mediated and linked to a specific HLA type. A variant allele in the *HLA-B* gene, a critical part of the immune system, was shown to strongly correlate with developing hypersensitivity. A number of papers, including the landmark PREDICT-1 study [28], showed that pharmacogenetic testing for *HLA-B\*57:01* resulted in a significantly decreased incidence of abacavir hypersensitivity reactions. Today, screening for *HLA-B\*57:01* is considered routine clinical practice before prescribing abacavir. Allele frequency of this HLA variant in different populations is given in Table 31.4 [29].

#### 31.5.3.2 HLA testing for prediction Stevens–Johnson syndrome with the use of the antiepileptic carbamazepine

CBZ is a key component of the pharmacotherapy for epilepsy, bipolar disorder, trigeminal neuralgia, and chronic pain. The drug is however associated with the risk of hypersensitivity reactions that range from benign conditions such as urticaria to life-threatening disorders, such as SJS and toxic epidermal necrolysis (TEN), which are associated with considerable mortality. The association between CBZ-induced SJS/TEN with HLA-1502 was first identified by Chung et al. [30]. The US-FDA subsequently made a labeling change to the product information related to CBZ recommending testing for HLA allele B\*1502 prior to treatment with CBZ and recommends genotyping in all Asians for this specific biomarker.

### 31.5.4 Testing for the activity of CYP2C19 prior to the use of clopidogrel

The antiplatelet drug clopidogrel is used in combination with aspirin for patients undergoing percutaneous interventions as well as for reducing the risk of myocardial infarction and stroke in patients with atherosclerotic vascular disease or established peripheral vascular disease. The drug is an inactive prodrug that gets converted to an active metabolite via CYP2C19. This metabolite irreversibly inhibits the platelet Adenosine diphosphate (ADP) receptor P2Y<sub>12</sub> to produce its effect. Individuals with a variant allele (especially the CYP2C19\*2 allele) are classified as poor metabolizers and cannot activate clopidogrel [31]. These individuals have been shown to have lower active metabolite levels, higher platelet reactivity, and poorer clinical outcomes. Approximately 2% of Caucasians, 4% of African-Americans, and 14% of Chinese are estimated to be CYP2C19 poor metabolizers. Both the Dutch Pharmacogenetics Working Group and the Clinical Pharmacogenetics Consortium recommend testing for CYP2C19 activity prior to initiating treatment with clopidogrel and using an alternate platelet inhibitor (prasugrel and ticagrelor) in poor metabolizers.

A representative list of 20 drugs from diverse therapeutic areas where preprescription pharmacogenetics/pharmacogenomics (use of a biomarker) is either recommended and/or also stated appropriately in varying sections of the product label is given in Table 31.5. This list has been adapted from the US-FDA website [32]. The entire list can be viewed on their website (<https://www.fda.gov/Drugs/ScienceResearch/ucm572698.htm>).

## 31.6 Validation of genetic/genomic information for drug response

While genetic or genomic information can be extremely valuable in improving medical decisions such as determining whether or not a patient is a good candidate for a particular drug, one should not underestimate the utility of other information such as age, gender, comorbidity, concomitant medication, or even patient preferences in choosing the right drug and dose. In other words, final treatment decisions should be holistic and not just based on the results of a genetic or genomics test alone. A case in point is that of the use of warfarin, described earlier in the chapter. The URL [www.warfarindosing.org](http://www.warfarindosing.org) [33] has a prediction tool/algorithm that makes use of multiple variables such as age, gender, ethnicity, race, height, weight, presence or absence of liver disease, and concomitant medications along with the results of genotyping for CYP2C9 and VKORC1 to arrive at the dosing requirement for warfarin in an individual patient. This is a good example that results of genomic testing alone are not sufficient to guide treatment decisions. Another example is the Framingham risk equation to estimate the risks of various types of cardiovascular outcomes among individuals at risk [34].

## 31.7 Clinical implementation of pharmacogenetics and pharmacogenomic information

Although pharmacogenomic/genetic-based research has been conducted for several years now, several challenges have slowed down its widespread implementation in clinical practice. Some of the barriers are discussed in the following sections.

### 31.7.1 Cost of testing and turnaround time

Just like any diagnostic test (e.g., peripheral smear for malaria), results of pharmacogenetic tests must be available as quickly as possible and preferably at low cost, particularly when treatment decisions must be made rapidly. If the costs are high, and a large number of patients are likely to test negative, there could be bravado (I will take my chances) on part of the clinician to prescribe a drug without genetic testing [35]. This includes even those drugs that have FDA labeling for pharmacogenetic testing such as warfarin that could cause life-threatening bleeding or deaths following use in patients with variant alleles (see earlier).

### 31.7.2 Limitations of single-nucleotide polymorphism testing in isolation

Testing for SNPs is useful only to when this can *entirely* predict the response to drug therapy. However, in reality, drug treatment, by and large is far more complex and is influenced not just by genetics, but also lifestyle, environment, and concomitant medications as well as the underlying disease condition and comorbidities such as liver and renal function. This is exemplified by warfarin, the action of which is influenced not only by the

**TABLE 31.5** Pharmacogenomic biomarkers in drug labeling with information about where matter is available in the product insert.

Drug	Therapeutic area	Biomarker	Product label information and sections where information is available
1. Abacavir	HIV/Infectious diseases	HLA-B	Dosage and administration, contraindications, warnings, and precautions
2. Amitriptyline	Psychiatry	CYP2D6	Precautions
3. Azathioprine	Rheumatology	TPMT	Dosage and administration, warnings, precautions, drug interactions, adverse reactions, clinical pharmacology
4. Boceprevir	Infectious diseases	IFNL3 (IL28B)	Clinical pharmacology
5. Brivaracetam	Neurology	CYP2C19	Clinical pharmacology
6. Clopidogrel	Cardiology	CYP2C19	Boxed warning, warnings and precautions, clinical pharmacology
7. Dapsone	Dermatology	G6PD	Warnings and precautions, use in specific populations
8. Darefenacin	Urology	CYP2D6	Clinical pharmacology
9. Dexlansoprazole	Gastroenterology	CYP2C19	Drug interactions, clinical pharmacology
10. Eliglustat	Inborn errors of metabolism	CYP2D6	Indications and usage, dosage and administration, contraindications, warnings and precautions, drug interactions, use in specific populations, clinical pharmacology, clinical studies
11. Flibanserin	Gynecology	CYP2C9	Clinical pharmacology
12. Glimepiride	Endocrinology	G6PD	Warnings and precautions, adverse reactions
13. Lidocaine and prilocaine	Anesthesiology	G6PD	Warnings and precautions, clinical pharmacology
14. Mycophenolic acid	Transplantation	HPRT1	Warnings and precautions
15. Ponatinib	Oncology	BCR–ABL1 (Philadelphia chromosome)	Indications and usage, warnings and precautions, adverse reactions, use in specific populations, clinical studies
16. Sodium nitrite	Toxicology	G6PD	Warnings and precautions
17. Warfarin	Cardiology/Hematology	CYP2C9 and VKORC1	Dosage and administration, drug interactions, clinical pharmacology
18. Warfarin	Hematology	PROS, PROC	Warnings and precautions
19. Ustekinumab	Dermatology and gastroenterology	IL12A, IL12B, IL23A	Warnings and precautions
20. Umeclidinium	Pulmonary	CYP2D6	Clinical pharmacology

HIV, Human immunodeficiency virus.

Adapted from the US FDA website (accessed 25.08.18).

genes previously mentioned but also by age, gender, concomitant medications, vitamin K intake among other factors. Predictions are also difficult when multiple genes contribute to the phenotype. In addition, testing for most human genes addresses testing only for the common variants, but not for the rare variants that can be easily missed by the testing.

### 31.7.3 Physician barriers

The average physician is not likely to have the awareness about the need to use pharmacogenetic testing or about the places where such testing is available or the cost. Of those who may be aware, there is likely to be insufficient knowledge, training, and experience [36] in using these tests for decision-making coupled with a genuine apprehension to actually incorporate them in the decision-making algorithm. The Clinical Pharmacogenetics Implementation Consortium (CPIC) was set up to address this need (see later).



31.7.4 *The need for specific protocols to guide decision-making post–pharmacogenetic testing*

Clinicians using pharmacogenetic testing need to have algorithms that are predefined so that the best use of the genetic information is made. These pharmacogenetic protocols must also be validated prior to clinical implementation.

31.7.5 *Lack of strong evidence/weak evidence base*

A fair amount of studies on genetic information applied to patient care are either retrospective, have small effect sizes differ in methodology or have significant heterogeneity [37]. Information from these studies needs to be validated in controlled trials to show that addition of genotyping to the existing standard of care actually improves patient outcomes. Prior to clinical implementation, large prospective hypothesis-driven studies in multiple ethnic groups are needed to truly determine if a particular pharmacogenomic testing protocol should add to the existing standard of care.

31.7.6 *Stakeholder engagement including the regulator*

A key aspect of translation of pharmacogenetic/pharmacogenomic testing is prompt review by the regulator and insistence on incorporating this information into the product label by the pharmaceutical company. A recent example from India is the regulator demanding a label change in the antiepileptic CBZ to include genetic testing for HLA \*1502, which predicts the risk of development of SJS.

31.8 Clinical pharmacogenetics implementation consortium—helping clinicians understand and apply pharmacogenetic information in practice

The advent of pharmacogenetics testing and its meteoric explosion led to the need for clinical guidelines on how best to understand and use pharmacogenetic tests, which can present a significant challenge to the practicing clinician. In 2009 the CPIC [38] was set up to address the gap between translating genetic testing done in the laboratory to actual patient care. The consortium consists of a limited number of dedicated staff and individual volunteers and posts periodically (and with regular updation), peer-reviewed, evidence-based, detailed gene/drug clinical practice guidelines. Over and above this, several countries around the world have their own research initiatives to help translate pharmacogenetic findings to practice, which include (1) the Pharmacogenomics Research Network in the United States, (2) the Canadian Pharmacogenomics Network for Drug Safety, (3) the International Warfarin Consortium (several countries), and (4) Pharmacogenetics Working Group of the Royal Dutch Association. Table 31.6 gives some guidance about using genetic testing in clinical practice.

TABLE 31.6 Points to be remembered while using pharmacogenetic testing for the practicing clinician.

1. There exists an evidence base for drug response based on genetic testing. Before prescribing a drug, check whether the drug requires genetic testing
2. Before prescribing any drug, check whether the product information recommends genetic testing. The US-FDA website gives comprehensive information for many drugs as to whether genetic testing is recommended in the product information
3. Assess the cost of genetic testing prior to recommending it and in particular when the patient has to pay for it
4. When the results of genetic testing for a drug are available, use web-based resources (e.g., the CPIC) to tailor drug therapy and monitor response based on the recommendations given
5. Genetics is just one component of bigger picture and more often than not, response to drugs is multifactorial. This needs to be borne in mind while treating an individual patient

CPIC, Clinical Pharmacogenetics Implementation Consortium; US-FDA, United States Food and Drug Administration.

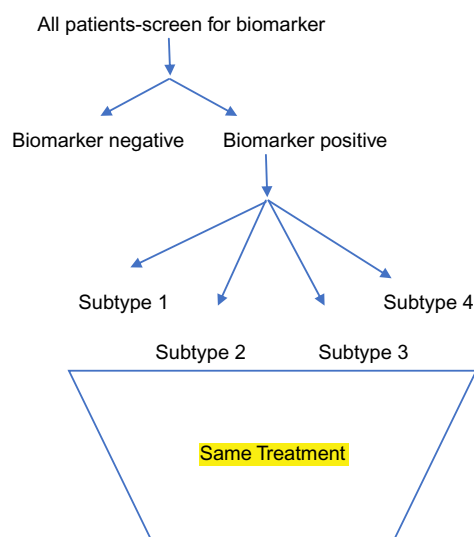
### 31.9 Pharmacogenomics and drug development—novel study designs in precision medicine

Pharmacogenomic studies are more often than not piggybacked on other pharmacological research studies though there do exist examples of prospective clinical trials, which are designed specially to test the impact of use of pharmacogenomics. Study designs in the area of oncology have taken center stage with the use of “basket” and “umbrella” trials, which are essentially biomarker-driven studies. The former trials are single treatment and single biomarker trials with different histologies being placed in baskets. The flowchart in Fig. 31.1 describes a typical basket (also called bucket) trial design.

An example of a basket trial is vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations [39]. In this Phase 2 study, patients with various nonmelanoma cancers were screened for the presence of BRAF V600 biomarker. The primary end point was the response rate, while the secondary end points were progression-free and overall survival. A total of 122 patients with the BRAF V600 mutations were enrolled. The “baskets” that were formed included NSCLC, Langerhans cell histiocytosis, pleomorphic xanthoastrocytoma, anaplastic thyroid cancer, cholangiocarcinoma, salivary-duct cancer, ovarian cancer, clear-cell sarcoma, and colorectal cancer. Preliminary vemurafenib activity was observed in NSCLC (42%) and Langerhans cell histiocytosis (43%).

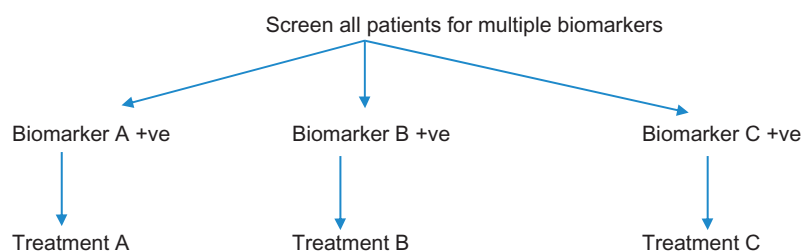
In contrast to basket trials, the umbrella trial evaluates many treatments within a single histology. Here, a multiplex assay (screening for multiple biomarkers) is used for determining the eligibility and treatment arm. The flowchart in Fig. 31.2 describes an umbrella trial.

The ALCHEMIST trial is an example of an umbrella trial for patients with resected nonsquamous, NSCLC. Tumor samples from patients are genotyped for EGFR mutations and ALK rearrangement. The presence of these mutations is a prerequisite for participation in the study. If positive for EGFR, patients are enrolled in the ALCHEMIST-EGFR study, a randomized, placebo-controlled trial of adjuvant erlotinib. Similarly, those positive for the ALK rearrangements go into the ALCHEMIST-ALK, a randomized placebo-controlled study of crizotinib [40,41]. The distinction between the two types of trials lies in the fact that umbrella trials take patients with the same type of cancer and assign them to different arms of a study based on their genomic information, while



**FIGURE 31.1** A schematic depiction of a basket trial design in precision medicine.

All patients are treated with the same treatment and then evaluate response based on subtype



**FIGURE 31.2** An schematic depiction of an umbrella trial in cancer precision medicine.

basket trials' group patients by mutation, regardless of which organs are involved. Both trials have the potential to dramatically alter the way NSCLC is treated.

Traditional randomized controlled study designs can also incorporate genotype-guided information. One such example is the GIFT randomized trial—effect of genotype-guided warfarin dosing on clinical events and anticoagulation control among patients undergoing hip or knee arthroplasty. The primary end point was the composite of major bleeding, an INR of 4 or greater, venous thromboembolism, or death. The study showed that genotype-guided warfarin dosing, compared with clinically guided dosing, significantly reduced the incidence primary end point in the genotype-guided dosing group [42].

### 31.10 Ethical issues in genomic medicine

The ethical concerns surrounding genomic testing can broadly be divided into three categories—equitable provision of health care, the possibility that genetic variants may get linked to race or ethnicity, and the subsequent consequences and the questions of consent, access, and privacy surrounding pharmacogenomic information [43]. As we increasingly move toward PM, clinical trials and patient-care protocols are now designed to collect genetic information. A large amount of DNA gets collected in the process and is more often than not, stored for posterity. The European Medicines Agency (EMA), US-FDA, and PMDA (Pharmaceutical and Medical Devices Agency, Japan) all strongly encourage prospective collection and storage of DNA samples from all consenting participants in clinical studies. This DNA will provide information not just relevant to the trial, but also beyond the trial. The two issues linked to this are that of consent and ownership of data. It is quite likely that participants who say yes to using their DNA for “future research” and give a “broad consent” (one that goes beyond the study being undertaken) are not fully aware of the implications of this consent. In clinical trials, because the research participant and the patient are usually one and the same person and the physician treating them is also the primary researcher, autonomy must be respected and preserved. Institutional review boards (IRB) have addressed this issue by asking for “reconsent” or at times a “tiered consent” [44]. The former requires that the investigator counsels and consents the research participant once more for a new study, while the latter consent is a process where participants can proactively specify what they permit (or otherwise) with regards to their DNA. This also places greater onus on IRBs and researchers to keep track of the consent forms signed by the participants and the directives therein.

The vast amounts of genetic information collected raise the obvious question of confidentiality and data ownership. Genetic information depending upon its nature may want to be kept confidential by the patient/participant. Yet, this information may be requested by family members with regard to understanding their genetic predisposition in relation to their own health. Insurance companies may seek this information for risk stratification and determination of cost of premiums to be paid. Employers may seek it for evaluating cost of hiring and provision of health insurance.

With regard to stored samples, there is the sticky issue of linking stored and anonymized samples with patient records. Anonymized genomic data will remain abstract unless correlated with clinical information and this presents potential breach of confidentiality. There is also the issue of associating “race” with genotype. A beneficial example that stands out is of that of susceptibility to CBZ-induced hypersensitivity. It is now known that diverse *HLA* alleles in different ethnic groups are associated with this particular adverse reaction. Han Chinese who carry the *HLA-B\*1502* allele are at greater risk of this adverse reaction. These individuals can easily be identified before prescription of CBZ and thus preempting and avoiding the use of the drugs. However, similar information can quite easily be misused depending upon the nature of testing. Finally, there is the issue of “distributive justice,” which mandates that the benefits and burdens of research be shared in an equitable manner. Depriving a certain ethnic group or groups of the benefits merely based on results of genetic testing is a serious issue that demands thought and attention.

### 31.11 Resources to collect and curate pharmacogenetic variants

PharmGKB ([www.pharmgkb.org](http://www.pharmgkb.org)) [45] is a comprehensive publicly available resource that collects and curates knowledge about the impact of human genetic variation on drug response. This organization manually curates associations between genetic variants and drug responses as reported in literature [46] and then adds them to the preexisting knowledge base such as the Human Genome Organization Gene Nomenclature Committee and the Human Genome Variation Committee. It is funded by the NIH and managed by the Stanford University.

The organization provides pharmacogenetic-based drug-dosing guidelines from CPIC, as well as the Royal Dutch Association for the Advancement of Pharmacy Pharmacogenetics Working Group and professional societies such as The American College of Rheumatology. PharmGKB also curates and annotates drug labels containing genetic information from both the US-FDA and the EMA.

### 31.12 The future of pharmacogenetics, pharmacogenomics, and personalized medicine

The genetic basis of variability in drug response can be broadly divided into three categories—(1) monogenic or the Mendelian traits, for example, the inherited disorders, and some rare adverse reactions (see earlier for abacavir hypersensitivity) typically influenced by single, rare coding variants; (2) oligogenic traits—those that are influenced by a limited number of genes; and (3) complex traits that are multifactorial. While the evaluation of monogenic traits is well entrenched and oligogenically improved, it is the complex traits that are the most difficult to address [47].

The PM concept though not entirely novel today is still tremendously exciting for all stakeholders concerned including the industry and the health-care industry. Its inherent strength lies in its tremendous potential to change the way therapy is provided by not just tailoring treatment, but also preventive strategies that involve a combination of genomics, epigenomics, proteomics, and the patient's clinical profile. Because this branch of medicine is proactive, rather than reactive, it offers the promise of better efficacy and/or effectiveness, fewer adverse reactions, improved quality of life, and perhaps lower costs of treatment in the long run. Its impact will however only be realized when all stakeholders contribute in equal measure—patients by participating in studies that involve genomic testing, regulators by careful review of information and changing product labels and providing the requisite policy revisions in a timely manner, information technology experts who must develop smart tools to both analyze the data and protect it, academic researchers who should uncover newer insights into existing diseases and populations, physicians who must learn to understand and apply this information for better patient care, manufacturers who must invest in new diagnostic tools at reasonable costs and low turnaround times, insurance companies who must explore new business models, and policy makers and guideline developers who must explore PM treatment protocols and algorithms. Only then will we move from the paradigm “one size fits all” to the right drug, with the right dose at the right time to the right patient’ paradigm for more effective health care.

## References

- [1] Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *J Am Med Assoc* 1998;279:1200–5.
- [2] Nebert DW. Pharmacogenetics and pharmacogenomics: why is this relevant to the clinical geneticist? *Clin Genet* 1999;56:345–7.
- [3] Motulsky A. From pharmacogenetics and ecogenetics to pharmacogenomics. *Med Secoli* 2002;14:683–705.
- [4] Vogenberg FR, Isaacson Barash C, Pursel M. Personalized medicine: Part 1: Evolution and development into theranostics. *Pharm Ther* 2010;35:560–76.
- [5] Pirmohamed M. Pharmacogenetics and pharmacogenomics. *Br J Clin Pharmacol* 2001;52(4):345–7.
- [6] Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. *N Engl J Med* 2011;364(12):1144–53.
- [7] U.S Food and Drug Administration [Internet]. E15 Pharmacogenomics Definitions and Sample Coding; April 2008. <<https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073162.pdf>> [accessed 23.08.18].
- [8] Naik RP, Haywood C. Sick cell trait diagnosis: clinical and social implications. *Hematol Am Soc Hematol Educ Program* 2015;2015(1):160–7.
- [9] Crick F. The double helix: a personal view. *Nature* 1974;248:766–71.
- [10] Sanger F, Coulson A. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 1975;94:441–8.
- [11] Hood L, Rowen L. The Human Genome Project: big science transforms biology and medicine. *Genome Med* 2013;5(9):79.
- [12] National Human Genome Research Institute [Internet]. The Human Genome Project. <<https://www.genome.gov/12011239/a-brief-history-of-the-human-genome-project/>> [accessed 25.08.18].
- [13] Collins FS, McKusick VA. Implications of the human genome project for medical science. *J Am Med Assoc* 2001;285(5):540–4.
- [14] National Institute of Health [Internet]. What is precision medicine? <<https://ghr.nlm.nih.gov/primer/precisionmedicine/definition>>; 2018 [accessed 23.08.18].
- [15] National Cancer Institute [Internet]. NCI Dictionary of Cancer Terms. <<https://www.cancer.gov/publications/dictionaries/cancer-terms/def/personalized-medicine>> [accessed 23.08.18].
- [16] Southern EM. DNA microarrays. History and overview. *Methods Mol Biol* 2001;170:1–15.
- [17] Politi K, Herbst RS. Lung cancer in the era of precision medicine. *Clin Cancer Res* 2015;21(10):2213–20. Available from: <https://doi.org/10.1158/1078-0432.CCR-14-2748>.

- [18] Liao B-C, Lin C-C, Shih J-Y, Yang JC-H. Treating patients with ALK-positive non-small cell lung cancer: latest evidence and management strategy. *Ther Adv Med Oncol* 2015;7(5):274–90.
- [19] Lambein K, Praet M, Forsyth R, Van den Broecke R, Braems G, Matthys B, et al. Relationship between pathological features, HER2 protein expression and HER2 and CEP17 copy number in breast cancer: biological and methodological considerations. *J Clin Pathol* 2011;64:200–7.
- [20] Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Herceptin Adjuvant (HERA) Trial Study Team. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005;353(16):1659–72.
- [21] Slamon D, Eiermann W, Robert N, et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med* 2011;365(14):1273–83.
- [22] Romond EH, Jeong J-H, Rastogi P, et al. Seven-year follow-up assessment of cardiac function in NSABP B-31, a randomized trial comparing doxorubicin and cyclophosphamide followed by paclitaxel (ACP) with ACP plus trastuzumab as adjuvant therapy for patients with node-positive, human epidermal growth factor receptor 2–positive breast cancer. *J Clin Oncol* 2012;30(31):3792–9.
- [23] Viani GA, Afonso SL, Stefano EJ, De Fendi LI, Soares FV. Adjuvant trastuzumab in the treatment of her-2-positive early breast cancer: a meta-analysis of published randomized trials. *BMC Cancer* 2007;7:153.
- [24] Kamali F, Khan TI, King BP, et al. Contribution of age, body size, and CYP2C9 genotype to anticoagulant response to warfarin. *Clin Pharmacol Ther* 2004;75:204–12.
- [25] Carlquist JF, Horne BD, Muhlestein JB. Genotypes of the cytochrome p450 isoform, CYP2C9, and the vitamin K epoxide reductase complex subunit 1 conjointly determine stable warfarin dose: a prospective study. *J Thromb Thrombolysis* 2006;22:191–7.
- [26] Biswas M, Bendkhale SR, Deshpande SP, et al. Association between genetic polymorphisms of CYP2C9 and VKORC1 and safety and efficacy of warfarin: results of a 5 years audit. *Indian Heart J* 2018;70(Suppl 3):S13–19.
- [27] U.S Food and Drug Administration [Internet]. FDA Online Label Repository. <[https://www.accessdata.fda.gov/drugsatfda\\_docs/label/2011/009218s107lbl.pdf](https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/009218s107lbl.pdf)> [accessed 23.08.18].
- [28] Ma JD, Lee KC, Kuo GM. HLA-B\*5701 testing to predict abacavir hypersensitivity. *PLoS Curr* 2010;2:RRN1203.
- [29] Martin MA, Klein TE, Dong BJ, Pirmohamed M, Haas DW, Kroetz DL. Clinical pharmacogenetics implementation consortium guidelines for HLA-B genotype and abacavir dosing. *Clin Pharmacol Ther* 2012;91:734–8.
- [30] Chung WH, Hung SI, Hong HS, et al. Medical genetics: a marker for Stevens–Johnson syndrome. *Nature* 2004;428:486.
- [31] O'Connor SA, Hulot JS, Silvain J, Cayla G, Montalescot G, Collet JP. Pharmacogenetics of clopidogrel. *Curr Pharm Des* 2012;18(33):5309–27.
- [32] U.S Food and Drug Administration [Internet]. Pharmacogenomic Biomarkers in Drug Labeling. <<https://www.fda.gov/downloads/Drugs/ScienceResearch/UCM578588.pdf>> [accessed 28.08.18].
- [33] The Warfarin Dose Refinement (DR) Collaboration [Internet]. Warfarin Dosing; <<http://www.warfarindosing.org/Source/Home.aspx>> [accessed 23.08.18].
- [34] Mannan H, Stevenson C, Peeters A, Walls H, McNeil J. Framingham risk prediction equations for incidence of cardiovascular disease using detailed measures for smoking. *Heart Int* 2010;5(2):e11.
- [35] Relling MV, Altman RB, Goetz MP, Evans WE. Clinical implementation of pharmacogenomics: overcoming genetic exceptionalism. *Lancet Oncol* 2010;11(6):507–9.
- [36] Klein ME, Parvez M Md, Shin GJ. Clinical implementation of pharmacogenomics for personalized precision medicine: barriers and solutions. *J Pharm Sci* 2017;106:2368–79.
- [37] Khoury MJ, Little J, Higgins J, Ioannidis JP, Gwinn M. Reporting of systematic reviews: the challenge of genetic association studies. *PLoS Med* 2007;4:e211.
- [38] The Clinical Pharmacogenetics Implementation Consortium [Internet]. <<https://cpicpgx.org/>> [accessed 23.08.18].
- [39] Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay JY, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. *N Engl J Med* 2015;373(8):726–36.
- [40] Cappuzzo F, Ciuleanu T, Stelmakh L, Cicen S, Szczésna A, Juhász E, et al. Erlotinib as maintenance treatment in advanced non-small-cell lung cancer: a multicentre, randomised, placebo-controlled phase 3 study. *Lancet Oncol* 2010;11:521–9.
- [41] Govindan R, Mandrekas SJ, Gerber DE, et al. ALCHEMIST trials: a golden opportunity to transform outcomes in early stage non-small cell lung cancer. *Clin Cancer Res* 2015;21(24):5439–44.
- [42] Do EJ, Lenzini P, Eby CS, Bass AR, McMillin GA, Stevens SM, et al. Genetics informatics trial (GIFT) of warfarin to prevent deep vein thrombosis (DVT): rationale and study design. *Pharmacogenomics J* 2012;12(5):417–24.
- [43] McKinnon RA, Ward MB, Sorich MJ. A critical analysis of barriers to the clinical implementation of pharmacogenomics. *Ther Clin Risk Manage* 2007;3(5):751–9.
- [44] The Indian Council of Medical Research [Internet]. National Ethical Guidelines for Biomedical and Health Research Involving Human Participants 2017; <[https://www.icmr.nic.in/sites/default/files/guidelines/ICMR\\_Ethical\\_Guidelines\\_2017.pdf](https://www.icmr.nic.in/sites/default/files/guidelines/ICMR_Ethical_Guidelines_2017.pdf)> [accessed 23.08.18].
- [45] The Pharmacogenomics Knowledgebase[Internet]. <<https://www.pharmgkb.org/>> [accessed 23.08.18].
- [46] Hewett M, Oliver DE, Rubin DL, Easton KL, Stuart JM, Altman RB, et al. PharmGKB: the pharmacogenetics knowledge base. *Nucleic Acids Res* 2002;30(1):163–5.
- [47] Zhang G, Nebert DW. Personalized medicine: Genetic risk prediction of drug response. *PharmacolTher* 2017;175:75–90. Available from: <https://doi.org/10.1016/j.pharmthera.2017.02.036>.

## Further reading

García-González J, Tansey KE, Hauser J, Henigsberg N, Maier W, Mors O, et al. Pharmacogenetics of antidepressant response: A polygenic approach. *ProgNeuropsychopharmacolBiol Psychiatry*. 2017 Apr 3; 75:128–34. Available from: <https://doi.org/10.1016/j.pnpbp.2017.01.011>.



# Integrated genomic and molecular medicine

Dhavendra Kumar<sup>1,2,3</sup>

<sup>1</sup>Institute of Medical Genetics, University Hospital of Wales, Cardiff, United Kingdom <sup>2</sup>Genomic Policy Unit, University of South Wales, Pontypridd, United Kingdom <sup>3</sup>William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

## 32.1 Introduction

Recent innovations and new developments in molecular biology, biotechnology, genomics, and many other omic sciences have revolutionized the contemporary and future practice of medicine. The diagnosis of most complicated and rare conditions is now possible with high degree of precision. The “gene-specific” and “genome-driven” diagnoses in many inherited and genetic disorders are now possible. The understanding of molecular mechanisms in a number of common and complex medical conditions has vastly improved. Progress in targeted genetic and molecular approach in pharmacotherapy has led many improvements from the current therapeutic regimens, which are designed for the “average model patient” and “one-size-fits-all” model approach. However, this has changed dramatically with rapid advances made in *evidence-based precision and personalized medicine* [1]. The whole process involves a stepwise approach in building the holistic picture referred to as stratified medicine with the ultimate aim of individualized or personalized therapeutic interventions [2]. The scope and limitations in the new exciting field of stratified and personalized medicine are reviewed in this chapter.

During the last decade, new genomic diagnostic tools and molecular innovations have led to the emergence of *precision medicine*, an innovative approach to disease prevention and treatment, which takes into account individual differences in people’s genes, environments, and lifestyles. The precision medicine is central to stratified and personalized medicine. It provides the clinician with tools to better understand the complex molecular mechanisms underlying a patient’s health, disease, or condition, and to better predict which treatments will be most effective. Advances in precision medicine have already led to powerful new discoveries and several new treatments, which are tailored to specific characteristics of individuals, such as a person’s genetic makeup, or the genetic profile of an individual’s tumor. This is leading to a transformation in the way we can treat diseases such as cancer. Patients with breast, lung, and colorectal cancers, as well as melanomas and leukemia, for instance, routinely undergo molecular testing as part of patient care, enabling physicians to select treatments that improve chances of survival and reduce exposure to adverse effects [3]. The potential for precision medicine to improve care and speed up the development of new treatments has only just begun to be exploited. Translating initial successes to a larger scale will require a coordinated and sustained global effort. Through collaborative public and private partnerships, the stratified and precision medicine initiatives will harness advances in genomics and biotechnology for accelerating biomedical discoveries [4].

## 32.2 Genetic, genomic, and molecular revolutions in medicine

Following the phenomenal discovery of the structure of the nucleic acids (DNA and RNA) and subsequent sequencing of the human genome, rapid progress is continually being made in understanding the genetic and molecular bases of human disease. In early stages the focus was largely on the peptide molecular basis, for

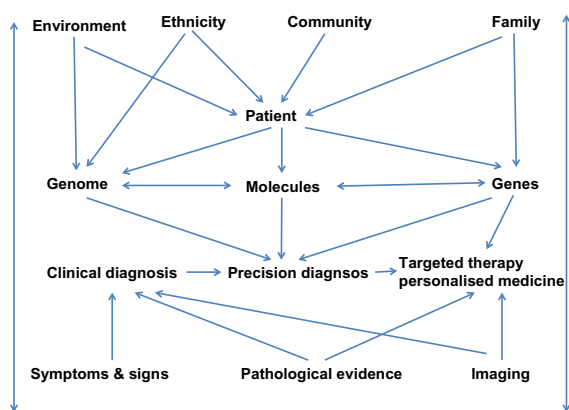
example, deciphering the structure of hemoglobin molecule to unravel the complexities of a number of inherited and acquired blood diseases. This period was the hallmark of molecular medicine. Nevertheless, developments in Mendelian genetics (single-gene diseases) facilitated understanding of the causation of human disease in the context of genes, inheritance patterns, recurrence risks, and genetic counseling [5]. However, these advances and evolving trends in medicine were restricted due to limited laboratory diagnosis. The practice of clinical molecular medicine in genetic and genomic terms is now a reality and most clinicians are emerging with new skills and competencies. This juncture changed dramatically with the sequencing of the human genome, opening new horizons for clinical medicine and even extending into domains of public and population health [6].

Although only less than 1% of the human genome sequence is different in any two individuals, variable nucleic-acid coding or noncoding sequences in the remainder of the genome could be functionally relevant to individual genetic constitution or genomic signature [7]. This “personalized sequence variation” is undeniably important and is agreeably the fundamental basis of genomic medicine. Functional annotation for individual sequence variation, when complete, will be crucial in precision diagnosis and personalized therapeutics [8]. This is likely to be vastly improved with the availability of targeted sequencing of selected genes, exons, or promoters. Currently, the focus is on whole-genome sequencing (WGS) that should reveal a full range of variants in both coding and noncoding genome. Variants that cause amino acid changes, and thus altered protein product, are in general dissimilar (nonsynonymous) compared to those that lack such an association (synonymous). If an excess of nonsynonymous substitution is observed for one particular coding region then this can be taken as an indicator of diversifying (positive) selection. With the help of next-generation sequencing technologies, more and more variants are being characterized and sequence annotations made available. It is envisaged that ultimately a fuller picture will emerge of the variants that alter genome function and will enable selection of those that contribute to health and disease in a particular individual [9].

During the last decade, rapid and unprecedented progress has been made in applied and translational genomic research leading to practical and dynamic utilizations in clinical medicine. Complicated laboratory techniques of genome sequencing (whole-exome, targeted deep-capture, and whole-genome) are no longer confined to research settings [10]. With the advent of next-generation sequencing the speed of generating enormous genome-sequencing data is considerably greater at successively lower costs [8]. An individual may get personal genome sequenced at around US\$500. Clinical diagnostic requests are now routinely made for array-comparative genomic hybridization, whole-exome/whole-genome sequencing by the broad range of specialists and general medical and health practitioners or even directly by the consumer [11,12]! Sincere efforts are put in place, both in public and private sectors, for establishing the role of genome sequencing in clinical medicine and public health. Higher state-level commitments are declared to set standards and guidelines for the genome sequencing in both clinical and research settings. The recent 100,000 genomes project of the UK government is a good example. The project, now aimed at 5 million genomes, offers a unique opportunity for bridging the gap between innovations and clinical applications, whether state or privately funded.

In the context of “precision and personalized medicine,” genetic, genomic, and molecular laboratory techniques are now being increasingly applied to select patients based on specific genetic and molecular signatures for a clear unambiguous diagnosis and selection of drugs pertinent to a particular therapeutic regimen. This approach requires properly validated scientific and clinical evidence at multiple levels within the agreed algorithm. The stepwise manner in the whole process has gained recognition and momentum in recent years leading to the emergence of “stratified medicine,” an umbrella term to encompass several groups of clinical scientists, physicians, and healthcare professionals. The core aim of these emerging concepts remains “precision diagnosis and management tailored to individual’s needs.” Thus fundamentally the *stratified medicine* offers the system for *precision medicine*, which implies that an individual patient’s clinical care is based on specific risk of disease or response to therapy by using diagnostic tests or techniques, whether conventional or genetic [13]. The whole model is set in the background of personal lifestyles, the structure and function of the family, sociocultural variation, ethnicity, and the community at large (Fig. 32.1). This model is beginning to yield dividends for patients and healthcare providers from targeted and effective treatments, whereas industry benefits from the potential for more efficient therapeutic developments as well as the market expansion for novel therapeutic drugs and devices.

The development of stratified medicine—based precision and personalized medicine is now pursued globally. To most medical and health practitioners the concept and philosophy behind “stratified medicine” are not unfamiliar. However, this approach is now remarkably strengthened with increasing accuracy and sophistication of the genomic and molecular medicine. It is widely acknowledged that the power of precision and personalized medicine is considerably enhanced with the availability of individual genome-sequence information [14].



**FIGURE 32.1** Emerging model of integrated genetic, genomic, and molecular medicine based on individual clinical evidence in the background of personal lifestyle, family history, ethnicity, sociocultural variation, and the community at large.

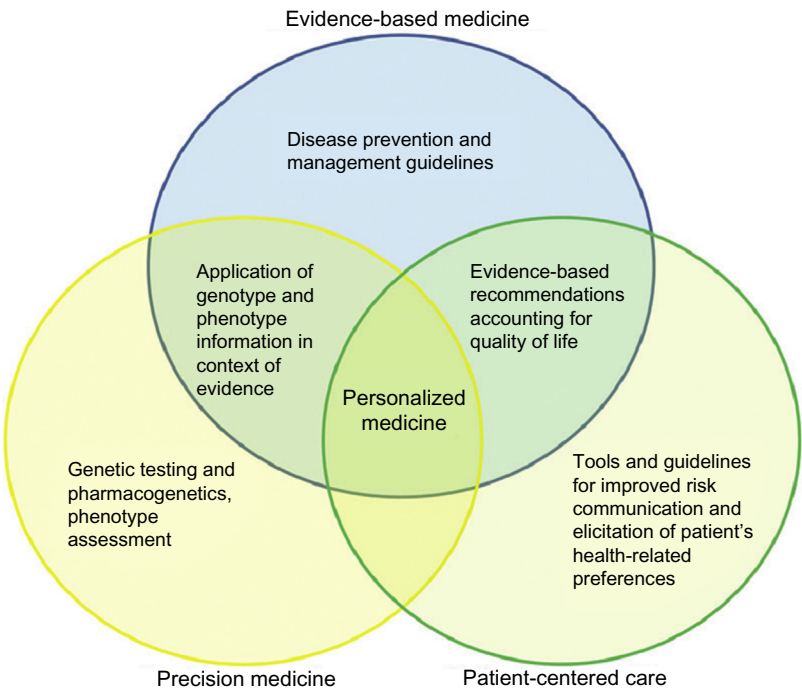
### 32.3 Evidence-based, precision, and personalized medicine

The success of genomic and molecular medicine will depend upon the ability to sequence an individual's full genome. With the benefit of new technologies, it is possible to generate gigabases of data as short sequence reads and to assemble the data accurately using the finished sequence as a template. This will provide the essential database of human genome variation for a given population. Comparison of these data sets will provide a full profile of common genome variation along each chromosome. Detection of each variant will help in estimating the recombination rates and correlation along each chromosome. This approach could give important baseline information on healthy tissue compared to pathological tissue. For example, a comparison of the cancer-genome sequences could allow monitoring the DNA changes on a genome-wide basis for cancer development. A similar approach could also be applied for other diseases. This genomic information on both healthy and diseased tissue could be used in screening an individual's disease risk and devising appropriate therapy and medical advice, paving the way forward for *personalized medicine* (Fig. 32.2; [15]).

As the human genome functional annotation becomes available, the prospects of "personalized medicine" will improve. A hypothetical scenario is described [16], where variation in the *PPAR- $\gamma$*  gene, one of the susceptibility genes in type 2 diabetes mellitus, is employed in selection of the most appropriate oral hypoglycemic drug.

The use of personal genetic information in a clinical setting could be requested and consented by the individual concerned. The individual sequence acquired could be restricted to one or two genotypes or as much as a complete genome sequence. The information thus acquired would be exclusive and private and wholly owned by the individual. It could be stored electronically, protected by a high-security code requiring unique personal identifiers, such as used for storing multiple fingerprint or iris pattern, for access only with the consent of the individual. The information might be taken either before consultation or afterwards and in either case would be subject to counseling by the medical practitioner and consent by the individual. The clinical consultation could initiate a specific investigation. The personal annotated genetic information, such as a set of gene mutations or variants for cardiovascular disease, of the individual patient would be made available for interpretation with respect to the clinical phenotype. The clinician would use the available risk information concerning each variant to provide a genetic assessment for the individual. In the case illustrated the individual has the heterozygous genotype TC at position 3:12,450,610. This corresponds to having both Pro 495 and Ala 495 forms of the protein *PPAR- $\gamma$* . This genotype confers an increased risk of insulin-resistant diabetes mellitus on the individual and also resistance to the thiazolidinedione class of antidiabetic drugs. Combining this with risk information for other genotypes would help to make informed subsequent clinical decisions. Thus with easy access to a well-annotated human genome and availability of cheap, accurate WGS technology, an individual could acquire either a specific or complete genetic-health profile, including risk and resistance factors. The information could then be used to improve and guide important medical decisions, to assess the risk of possible future exposures, and to select preventive treatments for improved health [16].

In brief the practice of personalized or specifically individualized medicine will become the central focus of the future practice of clinical medicine. However, this will demand a lot of commitment, perseverance,



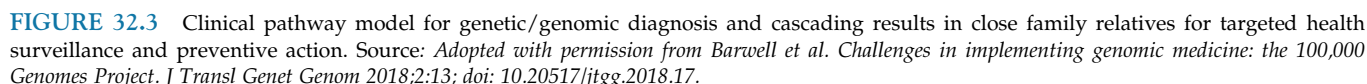
**FIGURE 32.2** Integrated relationship of personalized medicine with evidence-based medicine, precision medicine, and patient-centered health care. Source: Adopted from Kang SK et al. A road map of personalized care in radiology. Radiology 2015;277 (3):638–43.

and investment at personal, family, community, and public or state levels. Inevitably and understandably, this approach will raise several ethical and social concerns for the fear of inequity, discrimination (primarily due to enormous costs and affordability), and potential misuse or abuse (malpractice). The practice of personalized medicine shall not be allowed to develop without relevant professional and statutory safeguards put in place. This approach should be one of the other major ingredients of clinical practice pathway, what is often referred to “the 4 Ps of medicine”: medicine that will be more *predictive, personalized, preemptive, and participatory* [17,18]. Along with this list the new fifth term *precise* is being increasingly used (Table 32.1).

To reach these key long-term goals, National Institutes of Health (NIH-USA), National Institute of Health Research (NIHR-UK), and many other organizations are actively pursuing and promoting research in the aforementioned areas. These organizations are strategically investing in research to further our understanding of the fundamental causes of diseases at their earliest genetic, genomic, and molecular stages. The central theme of personalized medicine is based on the simple basic concept that individuals respond differently to environmental factors including therapeutic interventions, according to their genetic/genomic endowment and their own behavior and lifestyle. In the future, applied and translational genomic and molecular research will allow us to *predict* how, when, and in whom a disease will develop. We can envision a time when we will be able to *precisely* target or stratify treatment on a *personalized* (individualized) basis to those who need it, avoiding treatment to those who do not. Ultimately, this individualized approach will allow us to *preempt* disease before it occurs, utilizing the *participation* of individuals, communities, and healthcare providers in a *proactive* and *preparatory* fashion, as early as possible, and throughout the natural cycle of a disease process [18,19].

**TABLE 32.1** Acronyms of personalized genomic medicine and health care [17].

Genomics in medicine and health care
<ul style="list-style-type: none"><li>• Precision diagnosis</li><li>• Personalized information and counseling</li><li>• Personalized targeted therapy</li><li>• Preemptive approach</li><li>• Prediction and prevention of complications</li><li>• Preparatory and targeted planning for long-term health surveillance and care</li><li>• Participation in long-term management through lifestyle and behavior modifications</li></ul>



### 32.4 The stratified medicine

## II. Molecular medicine in clinical practice



TABLE 32.2 Criteria for stratified medicine (The Academy of Medical Sciences, United Kingdom) [19].

1. Continued research to understand the genetic and molecular bases of diseases
2. Development and use of increasingly sophisticated and powerful informatics technology
3. Improvement and standardization of clinical data collection and linkage with genomic and other databases
4. Increased collection of tissues for biomarker research and evaluation, and its organization in national and international biobanks
5. Greater efficiency and productivity in the development of therapeutics and diagnostics
6. The introduction of flexible and novel approaches for the regulatory assessments of innovative stratified medicine products
7. Improved flexibility in pricing for stratified medicine products—both for the diagnostic and for the associated therapy—to ensure cost-effectiveness for payers while encouraging innovation

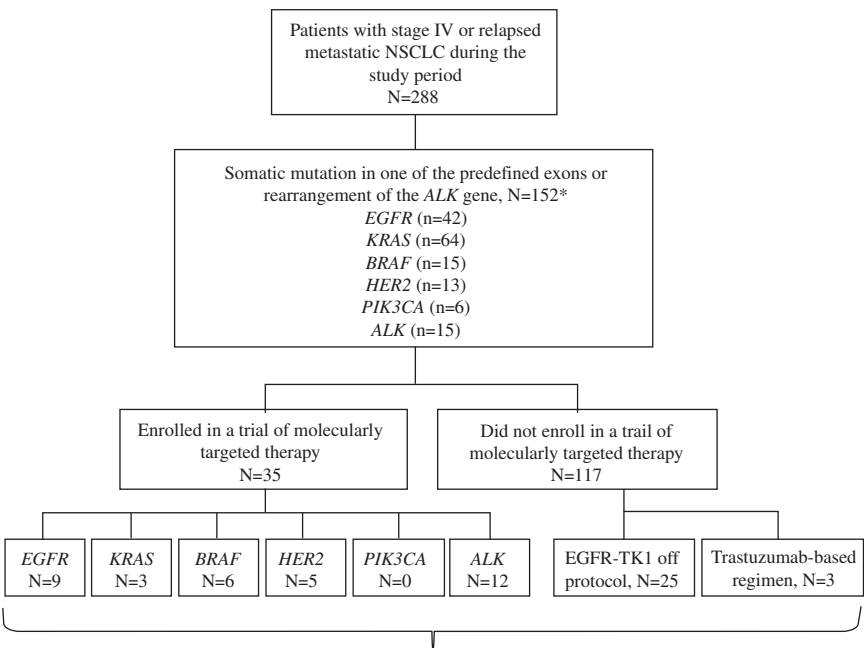


FIGURE 32.4 Flowchart of patients with stage IV or relapsed metastatic NSCLC onto molecularly targeted therapy during the study period [26]. NSCLS, Nonsmall cell lung cancer.

Several programs and incentives are now operational for “stratified medicine” to enable partnership across academia, industry, healthcare systems, regulatory/pricing authorities, research funders, and patient groups. The progress toward stratified medicine, increasingly confused with “personalized medicine,” relies fundamentally upon data, which is central to the applied and translational research to understand the molecular basis of disease; development of targeted interventions; effective regulation, health technology assessment, and valuation of stratified medicine products; and the stratification of treatment by physicians [25].

Among many examples of stratified approaches in planning and executing treatment for common cancers, the case for *nonsmall cell lung cancer* is noteworthy, probably the best paradigm in the context of stratified medicine (Fig. 32.4) [26]. Heterogeneity in patients, based on driver oncogenes mutations, is crucial for selecting targeted drugs for treatment [27].

There are several challenges and obstacles to realizing the full potential of benefits of the substantial progress in genomic and molecular research in pursuit of stratified approaches to clinical medicine:

- Standardization of genome-sequencing platforms to avoid laboratory-to-laboratory variability complicating the analysis of combined datasets.
- High levels of enrollment for sequencing are required to benefit from the accumulation of whole-genome sequence data, which will require that privacy and data-protection concerns be addressed.
- Because of the complexity, capital expense of equipment, and size of datasets, progress in molecular medicine is increasingly requiring collaboration between many academic groups, public institutions, and industry, often across countries.
- Genomic information on its own, although useful, is only part of the story. Greater knowledge is gained when such genetic information is linked to clinical outcomes. Thus there remains a major hurdle to link genome databases to healthcare records, which need to be electronic for this to be done efficiently.

- Research is still required so that genetic variations are not only correlated to diseases, but causal links are established, if the underlying molecular mechanisms of disease are to be understood.
- Correlation of genetic variation and disease may sometimes not transcend ethnic groups. The Pharmacogenetics for Every Nation Initiative has been set up to address this issue.
- The effect of epigenetic variations on drug response and pharmacoeugenomics needs further research [28]. Epigenetic variations are inheritable, affect gene expression levels and therefore phenotype, and yet do not result from changes in the DNA sequence [29].

### 32.5 Integrated genomic and molecular medicine

There are multiple factors that will determine the development and adoption of genomic and molecular approaches to medicine (Fig. 32.5). There are “pull” factors, in that the healthcare system needs to become increasingly effective and sustainable, in particular the economic policies for investment and cost reimbursement. There are also “push” factors, from recent advances in medical science and informatics; and the pharmaceutical industry requires substantial improvements in research and development productivity to remain a viable sector in the long term [30]. These factors accelerate the momentum of stratified medicine and are transformative in the provision of care. Detailed discussion on this aspect of stratified medicine is beyond the scope and remit of this chapter. However, the following major areas are important to consider for planners and developers of genomic medicine:

- Effective and sustainable healthcare systems
- Scientific and technological advances
- Diagnostic applications to accommodate new disease categories
- Challenges facing the pharmaceutical industry
- Role of the regulatory and statutory agencies

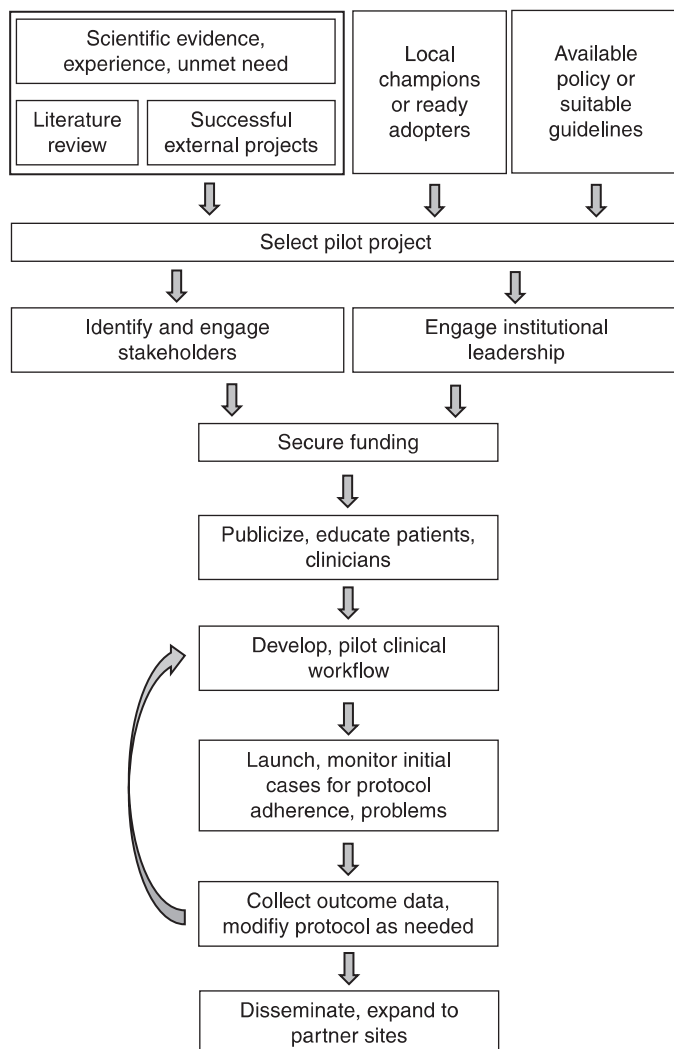
Successful implementation of genomic medicine strategy would depend on dealing with mammoth organizational challenges. The algorithm of a proposed clinical protocol and pathway requires multiagency and multispecialty approach (Fig. 32.5) [31]. All elements of phenotype recognition, clinical diagnosis, evidence-based genomic and molecular diagnosis, research efforts for annotating and recording all sequence variation information, parallel pilot projects, keeping the health care and public informed, and gaining support from relevant funding sources are all important.

Finally, the success of the UK 100,000 Genomes Project (Fig. 32.6) is noteworthy to ensure systemic evidence-based provision of genomic medicine services in the National Health Service (NHS). It is evident from the enthusiastic response from the UK government to expand the project to include 5 million genomes.

### 32.6 Summary

An important milestone in the history of medical science is the recent completion of the human genome sequence. The progress on identification of approximately 22,000 genes and their regulatory regions provides the framework for understanding the molecular basis of disease. This advance has also laid the foundation for a broad range of genomic tools that can be applied to medical science. These developments in gene and gene-product analysis across the whole genome have opened the way for targeted molecular genetic testing in a number of medical disorders. This is destined to change the practice of medicine: future medical practice will be more focused and individualized, often referred to as “personalized medicine.” However, despite these exciting advances, many practicing clinicians perceive the role of molecular genetics, in particular that of medical genomics, as confined to the research arena with limited clinical applications. Genomic medicine applies the knowledge and understanding of all genes and genetic variation in human disease. The basic ingredient of the contemporary practice of medicine is clinical molecular medicine that encompasses genetic, genomic, and molecular applications. This chapter introduces genomics-based advances in personalized disease-susceptibility screening, diagnosis, prognostication, stratified approach for genomics-led therapeutics, and prediction of treatment outcome in various areas of medicine.

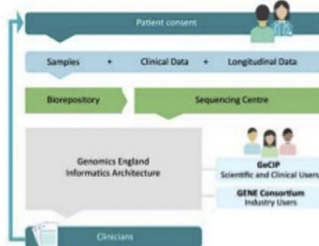
Finally, the art and science of the practice of medicine at all times are true reflections of dynamic adjustment of the physical state of the human body and environmental pressures. In this context the innate characteristics



**FIGURE 32.5** Landscape of integrated genomic and molecular medicine with major components of infrastructure development, research and development, service programs development, audit and quality control, and outcome measures. Source: *Courtesy of Manolio et al. Implementing genomic medicine in the clinic: the future is here. Genet Med 2013;15(4):258–67.*

## The 100,000 Genomes Project

- Launched by former **PM David Cameron** in 2012
- Delivered through a **partnership** between the NHS, Genomics England, NIHR, Public Health England, Health Education England
- Whole genome sequencing NHS patients with **rare disease, cancer and infectious disease**
- Recruitment through **13 NHS Genomic Medicine Centres** – recruiting through routine care, treating through routine channels
- **Involvement of all countries of UK**



**FIGURE 32.6** The 100,000 Genomes Project of United Kingdom ([www.genomicsengland.co.uk](http://www.genomicsengland.co.uk); [21]).

conferred by the genetic and genomic constitution provide the framework on which a range of lifetime environmental experiences and pressures would act and manifest in either positive or morbid (disease) states. This was echoed clearly over 100 years ago across the medical community in one of the classic Harveian Orations of the Royal College of Physicians in London, England—

It was in Padua that medicine, long degraded and disguised, was now to prove her lineage as the mother of natural science, and the truth of the saying of Hippocrates, that to know the nature of man one must know the nature of all things. *Sir Clifford Allbutt, Regius Professor of Physic, Harveian Oration (1900) [32].*

## References

- [1] Kumar D. The personalised medicine. A paradigm of evidence-based medicine. *Ann Ist Super Sanità* 2011;47(1):31–40. Available from: [https://doi.org/10.4415/aNN\\_11\\_01\\_08](https://doi.org/10.4415/aNN_11_01_08).
- [2] Matthews PM, et al. The emerging agenda of stratified medicine in neurology. *Nat Rev Neurol* 2014;10:15–26.
- [3] Hall IP. Stratified medicine: drugs meet genetics. *Eur Respir Rev* 2013;22:53–7. Available from: <https://doi.org/10.1183/09059180.00008312>.
- [4] Tuff-Lacey A, et al. A collaborative approach to enabling stratified cancer medicine in the UK. *Drug Discovery Today* 2015;20(12):1414–18. Available from: <https://doi.org/10.1016/j.drudis.2015.07.001>.
- [5] Kumar D. Clinical medicine in the genome era: an introduction. *Genom Clin Med* 2008;53:145.
- [6] World Health Organization, Advisory Committee on Health Research. Genomics and World Health: Report of the Advisory Committee on Health Research. World Health Organization; 2002.
- [7] Cooper DN, et al. Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum Mutat* 2010;31(6):631–55.
- [8] Hamburg MA, Collins FS. The path to personalized medicine. *New Engl J Med* 2010;363(4):301–4.
- [9] Green ED, Guyer MS. Charting a course for genomic medicine from base pairs to bedside. *Nature* 2011;470(7333):204–13.
- [10] Snijders AM, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001;29(3):263–4.
- [11] Berg JS, Khoury MJ, Evans JP. Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time. *Genet Med* 2011;13(6):499–504.
- [12] Bloss CS, Schork NJ, Topol EJ. Effect of direct-to-consumer genomewide profiling to assess disease risk. *New Engl J Med* 2011;364(6):524–34.
- [13] Bieber, T., Stratified medicine: a new challenge for academia, industry, regulators and patients. *Stratified Med*, 2013; p. 3.
- [14] Ziegler A, et al. Personalized medicine using DNA biomarkers: a review. *Hum Genet* 2012;131(10):1627–38.
- [15] Kang SK, et al. A road map of personalized care in radiology. *Radiology* 2015;277(Number 3):638–43.
- [16] Bentley DR. Genomes for medicine. *Nature* 2004;429(6990):440–5.
- [17] Hood LE, Galas DJ. P4 medicine: personalized, predictive, preventive, participatory a change of view that changes everything. *N, Biotechnol* 29, 2009, 613–624.
- [18] Xu LH, et al., The re-emerging concept of personalized healthcare. *Personalized Medicine* 5(5), 2008. Available from: <https://doi.org/10.2217/17410541.5.5.457>.
- [19] Yang YT, Wiley E, Leppard J. Individualized medicine and pharmacogenomics: ethical, legal and policy challenges. *J Med Person* 2011;9(2):48–57.
- [20] Richards, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–23.
- [21] Barwell, et al. Challenges in implementing genomic medicine: the 100,000 Genomes Project. *J Transl Genet Genom* 2018;2:13. Available from: <https://doi.org/10.20517/jtgg.2018.17>.
- [22] Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nat Rev Drug Discovery* 2007;6(4):287–93.
- [23] Silver MA, et al. BNP Consensus Panel 2004: a clinical approach for the diagnostic, prognostic, screening, treatment monitoring, and therapeutic roles of natriuretic peptides in cardiovascular diseases. *Congest Heart Fail* 2004;10(s5):1–30.
- [24] (UK), T.A.o.M.S. Realizing the potential of Stratified Medicine. 2013.
- [25] Khoury MJ, et al. Multilevel research and the challenges of implementing genomic medicine. *JNCI Monogr* 2012;44:112–20.
- [26] Cardarella S, Ortiz TM, Joshi VA, Butaney M, Jackman DM, Kwiatkowski DJ, et al. The introduction of systematic genomic testing for patients with non-small-cell lung cancer. *J. Thorac. Oncol.* 2012;7(12):1767–74.
- [27] Tsuboi M, et al. The present status of postoperative adjuvant chemotherapy for completely resected non-small cell lung cancer. *Ann Thorac Cardiovasc Surg* 2007;13(2):73.
- [28] Baer-Dubowska W, Majchrzak-Celińska A, Cichocki M. Pharmacoeugenetics: a new approach to predicting individual drug responses and targeting new drugs. *Pharmacol Rep* 2011;63(2):293–304.
- [29] Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;33:245–54.
- [30] Tambuyzer E. Towards a framework for personalized healthcare: lessons learned from the field of rare diseases. *Pers Med* 2010;7(5):569–86.
- [31] Manolio, et al. Implementing genomic medicine in the clinic: the future is here. *Genet Med* 2013;15(4):258–67.
- [32] Allbutt TC. The Harveian oration on physiological darkness before Harvey: delivered before the Royal College of Physicians on October 18th. *Br Med J* 1900;2(2078):1271.





# Glossary—molecular medicine\*

---

*Compiled and edited by Dhavendra Kumar*

- Acrocentric** a chromosome having the centromere close to one end, for example, chromosomes 21 and 22.
- Adaptive immune response** a response highly specific for a given pathogen that prevents its spread, commonly through cellular or humoral immunity producing protein-specific antibodies.
- Algorithm** a step-by-step method for solving a computational problem or a set of precise rules or procedures programmed into a pacemaker or defibrillator that is designed to solve a specific clinical problem.
- Allele** an alternative form of a gene at the same chromosomal locus.
- Allelic heterogeneity** different alleles for one gene, for example, CFTR, the gene for cystic fibrosis, has around 1000 different alleles.
- Allogenic** the transplantation of cells, tissues, and organs from one individual to another belonging to the same species.
- Alternative splicing** a regulatory mechanism by which variations in the incorporation of coding regions (see *exon*) of the gene into messenger RNA (mRNA) lead to the production of more than one related protein or isoform.
- Amino acid** a chemical subunit of a protein. Amino acids polymerize to form linear chains linked by peptide bonds called polypeptides. All proteins are made from 20 naturally occurring amino acids.
- Amplification refractory mutation system (ARMS)** an allele-specific PCR amplification reaction, for example, the genetic testing kit for cystic fibrosis.
- Annealing** the association of complementary DNA (or RNA) strand to form the double-stranded structure.
- Annotation** the descriptive text that accompanies a gene sequence in a database method commonly with reference to a particular phenotype.
- Antibody** a protein produced by the immune system in response to an antigen (see *antigen*). Antibodies bind to their target antigen to help the immune system destroy the foreign entity.
- Anticipation** a phenomenon in which the age of onset of a disorder is reduced and/or severity of the phenotype is increased in successive generations, for example, early age at onset with clinical severity of myotonic dystrophy in successive generations when inherited from an affected mother.
- Antigen** a molecule that is perceived by the immune system to be foreign and triggers the antibody formation.
- Apoptosis** programmed cell death or senescence, commonly used in reference to carcinogenesis.
- Aptamer** usually a single stranded oligonucleotide that binds to a target molecule, for example, a protein, by steric interactions.
- Arrhythmia** any heart rhythm that falls outside the accepted norms with respect to rate, regularity, or sequence of depolarization. (Any abnormal or absent heart rhythm.)
- Atrial tachycardia (AT)** a rapid heart rate that starts in the atria and includes atrial fibrillation (AF) and atrial flutter (AFL).
- Atrioventricular (AV) node** a section of specialized neuromuscular cells that are part of the normal conduction pathway between the atria and the ventricles. (A junction that conducts electrical impulses from the atria to the ventricles of the heart.)
- Atrioventricular (AV) synchrony** the normal activation sequence of the heart in which the atria contract and then, after a brief delay, the ventricles contract. The loss of AV synchrony can have significant hemodynamic effects. Dual chamber pacemakers are designed to attempt to maintain AV synchrony.
- Atrium** the heart is divided into four chambers. Each of the two upper chambers is called an atrium. (Atria is the plural form of atrium.) Either of the two upper chambers of the heart, above the ventricles that receive blood from the veins and communicate with the ventricles through the tricuspid (right) or mitral (left) valve.
- Autologous** when cells, tissues, and organs are grafted into the same individual from whom it was derived.
- Autosome** any chromosome other than a sex chromosome (X or Y) and the mitochondrial chromosome.
- Autozygosity** in an inbred person, homozygosity for alleles is identical by descent, commonly used in reference to consanguineous relationship for assessing the degree of relationship. *Autozygosity mapping* is a form of genotyping for autosomal recessive disorders in which affected individuals are expected to have two identical disease alleles by descent.
- Bacterial artificial chromosome (BAC)** DNA vectors into which large DNA fragments can be inserted and cloned in a bacterial host.
- Bioinformatics** an applied computational system which includes development and utilization of facilities to store, analyze, and interpret biological data; specific application for clinical use is commonly referred as clinical informatics.
- Biotechnology** the industrial application of biological processes, particularly recombinant DNA technology and genetic engineering, for example, insulin and hepatitis vaccines.

\* Disclaimer: Contents of this section of the book are derived from many public databases and published resources. The author/editor does not make any claim of ownership and allows its free use by any reader or user of this book.

- BLAST (basic local alignment search tool)** a fast database similarity search tool used by the NCBI that allows the world to search query sequences against the GeneBank database over the web; the PubMed or Medline is an excellent example.
- Blastocyst** the mammalian embryo at the stage at which it is implanted into the wall of the uterus.
- Bradycardia (Bradyarrhythmia)** a heart rate that is abnormally slow; commonly defined as under 60 beats per minute or a rate that is too slow to physiologically support a person and their activities.
- CABG** coronary artery bypass graft, the surgical intervention for the management of coronary artery diseases (CAD).
- Candidate gene** any gene or set of DNA sequences, by virtue of a known property (function, expression pattern, chromosomal location, structural motif, etc.), is considered as a possible locus for a given disease.
- Cardiac arrest** failure of the heart to pump blood through the body. If left untreated, it is dangerous and life threatening, often with fatal outcome.
- Cardioversion** termination of an atrial or ventricular tachyarrhythmia (other than ventricular fibrillation) by a delivery of a direct low energy electrical current synchronized to a specific instant during the cardiac rhythm.
- Carrier** a person who carries an allele for a recessive disease (see *heterozygote*) without the disease phenotype but can pass it on to the next generation. *Carrier testing* is carried out to determine whether an individual carries one copy of an altered gene for a particular recessive disease.
- Caspase** a cysteinyl aspartate-specific protease that hydrolyzes cellular proteins; caspases are the main mediators of apoptotic processes.
- CD—cluster of differentiation** molecules located on the cell surface which are commonly used to characterize and define immune cells.
- cDNA** (complementary DNA)—a piece of DNA copied in vitro from mRNA by a reverse transcription enzyme.
- Cell cycle** series of tightly regulated steps that a cell goes through from its creation to division to form two daughter cells.
- CentiMorgan (cM)** a unit of genetic distance equivalent to 1% probability of recombination during meiosis. One centiMorgan is equivalent, on average, to a physical distance of approximately 1 Mb in the human genome.
- “Central dogma”** a term proposed by the Nobel Laureate Francis Crick in 1957—*DNA is transcribed into RNA which is translated into protein.*
- Centromere** the constricted region near the center of a chromosome that has critical role in cell division.
- CHD** coronary heart disease.
- CHF** congestive heart failure.
- Chimera** a hybrid, particularly a synthetic DNA molecule that is the result of ligation of DNA fragments that come from different organisms or an organism derived from more than one zygote.
- Chromosome** subcellular structures which contain and convey the genetic material of an organism.
- Chromosome painting** fluorescent labeling of whole or part of a chromosome by the fluorescent in situ hybridization (FISH) procedure in which labeled probes each consist of complex mixture of different DNA sequences from a single chromosome.
- Class I antiarrhythmic drugs** drugs which act selectively to depress fast sodium channels, slowing conduction in all parts of the heart, for example, *quinidine*, *procainamide*, *flecainide*, *encainide*, and *propafenone*.
- Class II antiarrhythmic drugs** drugs which act as beta-adrenergic blocking agents, for example, *propranolol*, *metoprolol*, and *atenolol*.
- Class III antiarrhythmic drugs** drugs which act directly on cardiac cell membrane, prolong repolarization and refractory periods, increase VF threshold, and act on peripheral smooth muscle to decrease peripheral resistance, for example, *amiodarone* and *sotalol*.
- Clinical sensitivity** the proportion of persons with a disease phenotype who test positive.
- Clinical specificity** the proportion of persons without a disease phenotype who test negative.
- Clone** a line of cells derived from a single cell and therefore carrying identical genetic material.
- Cloning vector** a DNA construct, such as a plasmid, modified viral genome (bacteriophage or phage), or artificial chromosome, that can be used to carry a gene or fragment of DNA for purposes of cloning (e.g., a bacterial, yeast or mammalian cell).
- Coagulation factors** various components of the blood coagulation system. The following factors have been identified (synonyms which are or have been in use are included): Factor I (fibrinogen); Factor II (prothrombin); Factor III (thromboplastin, tissue factor); Factor IV (calcium); Factor V (labile factor); Factor VII (stable factor); Factor VIII [antihemophilic globulin (AHF), antihemophilic globulin (AHG), antihemophilic factor A Factor VIII C]; Factor IX [plasma thromboplastin component (PTC), Christmas factor, antihemophilic factor B]; Factor X (Stuart factor, Prower factor, Stuart–Prower factor); Factor XI [plasma thromboplastin antecedent (PTA), antihemophilic factor C]; Factor XII (Hageman factor, surface factor, contact factor); Factor XIII [fibrin stabilizing factor (FSF), fibrin stabilizing enzyme, fibrinase]; other factors [such as prekallikrein (Fletcher factor) and high molecular weight kininogen (Fitzgerald)].
- Coding DNA (sequence)** the portion of a gene that is transcribed into mRNA.
- Codon** a three-base sequence of DNA or RNA that specifies a single amino acid.
- Comparative genome hybridization (CGH)** use of competitive fluorescence in situ hybridization to detect chromosomal regions that are amplified or deleted, especially in tumors.
- Complementary DNA (cDNA)** DNA generated from an expressed messenger RNA through a process known as reverse transcription.
- Complex diseases** diseases characterized by risk to relatives of an affected individual which is greater than the incidence of the disorder in the population, for example, epilepsy, bronchial asthma, psoriasis, coronary artery disease, essential hypertension, and many other common diseases.
- Complex trait** one which is not strictly *Mendelian* (dominant, recessive, or sex linked) and may involve the interaction of two or more genes to produce a phenotype or may involve *gene–environment* interactions.
- Compound heterozygosity** two different heterozygous recessive alleles at a particular locus that cause genetic disease, for example, Hurler–Scheie mucopolysaccharidoses.
- Computational therapeutics** an emerging biomedical field concerned with the development of techniques for using software to collect, manipulate, and link biological and medical data from diverse sources. It also includes the use of such information in simulation models to make predictions or therapeutically relevant discoveries or advances.
- Computer-aided diagnosis (CAD)** a general term used for a variety of artificial intelligence techniques applied to medical images. CAD methods are being rapidly developed at several academic and industry sites, particularly for large-scale breast, lung, and colon cancer screening studies. X-ray imaging for breast, lung, and colon cancer screening are good physical and clinical models for the development of CAD

- methods, related image database resources, and the development of common metrics and methods for evaluation. [*Large-scale screening applications include (1) improving the sensitivity of cancer detection, (2) reducing observer variation in image interpretation, (3) increasing the efficiency of reading large image arrays, (4) improving efficiency of screening by identifying suspect lesions or identifying normal images, and (5) facilitating remote reading by experts (e.g., telemammography).*]
- Concordance** a measure of probability that two genetically identical individuals (twins) will share the similar physical or phenotypic resemblance; conventionally identical twins are close to 100% concordant compared to 50% concordance for nonidentical twins; measure of concordance is one of the accepted method of twin-based genetic association analysis in complex/common diseases.
- Congenital** any trait, condition, or disorder that exists from birth.
- Consanguinity** marriage between two individuals having common ancestral parents, commonly between first cousins; an approved practice in some communities who share social, cultural, and religious beliefs. In genetic terms, two such individuals could be heterozygous by descent for an allele expressed as *coefficient of relationship*, and any offspring could be therefore homozygous by descent for the same allele expressed as *coefficient of inbreeding*.
- Conserved sequence** a base sequence in a DNA molecule (or an amino acid sequence in a protein) that has remained essentially unchanged throughout evolution, for example, a number of benign (clinically insignificant) DNA polymorphic variants distributed across the genome.
- Constitutional mutation** an inherited mutation present in all cells containing the relevant nucleic acid sequences; by definition, it is similar to *germ-line mutation*.
- Contig** a consensus sequence generated from a set of overlapping sequence fragments that represent a large piece of DNA, usually a genomic region from a particular chromosome.
- Copy number** the number of different copies of a particular DNA sequence in a genome.
- Copy number variation (CNV)** variation in copy number sequences, likely to be of pathogenic importance for certain complex disease traits or rare malformation syndromes.
- CpG island** short stretch of DNA, often less than 1 kb, containing CpG dinucleotides that are unmethylated and present at the expected frequency. CpG islands often occur at transcriptionally active DNA; methylated CpG islands are important for epigenetic/epigenomic regulation.
- Cytoplasm** the internal matrix of a cell. The cytoplasm is the area between the outer periphery of a cell (the cell membrane) and the nucleus (in a eukaryotic cell).
- Defibrillation** termination of an erratic, life-threatening arrhythmia of the ventricles by a high energy, direct current delivered asynchronously to the cardiac tissue. The defibrillation discharge will often restore the heart's normal rhythm.
- Demographic transition** the change in the society from extreme poverty to a stronger economy, often associated with a transition in the pattern of diseases from malnutrition and infection to the intractable conditions of middle and old age, for example, cardiovascular disease, diabetes, and cancer.
- Dicer** an endoribonuclease belonging to the RNase III family that cleaves longer double-stranded RNAs and premicroRNAs into short double-stranded RNAs and mature microRNAs, respectively.
- Diploid** a genome (the total DNA content contained in each cell) that consists of two homologous copies of each chromosome; two haploid sets of 23 chromosomes make the diploid genome.
- Disease** a fluid concept influenced by societal and cultural attitudes that change with time and in response to new scientific and medical discoveries. In genomic context, similar collection of symptoms and signs (*phenotype*) may have very different underlying genomic sequences (*genotype*).
- Disease etiology** any factor or series of related events directly or indirectly causing a disease; for example, next-generation genome sequencing and proteomic techniques offer deeper insight of molecular mechanisms and biological processes.
- Disease expression** when a pathogenic genotype is manifested in the *phenotype*.
- Disease management** a continuous, coordinated health-care process that seeks to manage and improve the health status of a patient over the entire course of a disease. Disease management services include disease prevention efforts and as well as patient management involving a number of specialist and general health professionals, the multidisciplinary health care.
- Disease phenotype** includes disease-related changes in tissues as judged by gross anatomical, radiological, histological, and molecular pathological changes. Gene and protein expression analysis and interpretation studies, particularly at the whole genome or proteome levels, are capable of distinguishing apparently similar phenotypes.
- Diversity, genomic** the number of base differences between two genomes divided by the genome size.
- DNA (deoxyribonucleic acid)** the chemical that comprises the genetic material of all cellular organisms.
- DNA fingerprinting** use of hypervariable minisatellite probe (usually those developed by Sir Alec Jeffreys' probes) on a Southern blot to produce an individual-specific series of bands for identification of individuals or relationships. This is now much refined using the conventional and next-generation genome sequencing techniques.
- DNA sequencing** technologies through which the order of base pairs in a DNA molecule can be determined, for example, clinical exome or whole genome sequencing.
- Domain** a discrete portion of a protein with its own function. The combination of domains in a single protein determines its overall function. A particular domain in the peptide chain can be assigned to specific coding section (exon) of the gene.
- Dominant** an allele (or the trait encoded by that allele) which produces its characteristic phenotype when present in the heterozygous form.
- Dominant negative mutation** a mutation which results in a mutant gene product which can inhibit the function of the wild-type gene product in heterozygotes. In contrast, the gain of function-dominant mutation results in a novel or enhanced function of the gene.
- Dosage effect** the number of copies of a gene; variation in the number of copies can result in aberrant gene expression or associated with disease phenotype.
- Drug design** development of new classes of medicines based on a reasoned approach using gene sequence and protein structure function information rather than the traditional trial-and-error method.
- Drug interactions** refer to adverse drug interaction, drug–drug interaction, drug–laboratory interaction, drug–food interaction, etc. It is defined as an action of a drug on the effectiveness or toxicity of another drug.

- Ejection fraction** a measure of the output of the heart with each heartbeat (stroke volume divided by end-diastolic volume).
- Electrocardiogram (ECG)** a printout from an electrocardiography machine used to measure and record the electrical activity of the heart.
- Electronic health record (EHR)** a real-time patient health record with access to evidence-based decision support tools that can be used to aid clinicians in decision-making, automating, and streamlining clinician's workflow, ensuring that all clinical information is communicated. It can also support the collection of data for uses other than clinical care, such as billing, quality management, outcome reporting, and public health disease surveillance and reporting.
- Embryonic stem cells (ES cells)** a cell line derived from undifferentiated, pluripotent cells from the embryo.
- ENTREZ** an online search and retrieval system that integrates information from databases at NCBI. These databases include nucleotide sequences, protein sequences, macromolecular structures, whole genomes, OMIM, and MEDLINE, through PubMed.
- Environmental factors** may include chemical, dietary factors, infectious agents, physical, and social factors.
- Enzyme** a protein which acts as a biological catalyst that controls the rate of a biochemical reaction within a cell without disrupting its structure.
- Epigenetic** a term describing nonmutational phenomenon, such as methylation and histone modification, that modify the expression of a gene.
- Epigenetics** the study of heritable changes in gene activity that are not caused by changes in the DNA sequence; for example, changes in DNA methylation or histone modification alter gene expression without changing the underlying DNA sequence.
- Euchromatin** relatively lightly packed chromatin that is transcriptionally active.
- Eukaryote** an organism whose cells show internal compartmentalization in the form of membrane-bounded organelles (includes animals, plants, fungi, and algae).
- Exon** the sections of a gene that code for all of its functional product. Eukaryotic genes may contain many exons interspersed with noncoding introns. An exon is represented in the mature mRNA product—the portions of an mRNA molecule that is left after all introns are spliced out, which serves as a template for protein synthesis.
- Expressivity** variations in a particular phenotype among individuals carrying the same genotype; for example, variations in clinical severity of a disease phenotype in individuals carrying the same mutant allele for a Mendelian (single gene) disease.
- Family history** an essential tool in clinical genetics. Interpreting the family history can be complicated by many factors, including small families, incomplete or erroneous family histories, consanguinity, variable penetrance, and the current lack of real understanding of the multiple genes involved in polygenic (complex) diseases.
- Fluorescence in situ hybridization (FISH)** a form of chromosome in situ hybridization in which nucleic acid probe is labeled by incorporation of a *fluorophore*, a chemical group that fluoresces when exposed to UV irradiation.
- Founder effect** changes in allelic frequencies that occur when a small group is separated from a large population and establishes in a new location. In contrast the *founder mutation* refers to a specific mutation in a particular gene present in an ethnic migrant population that is prevalent in the indigenous population.
- Frame-shift mutation** the addition or deletion of a number of DNA sequences that is not a multiple of three, thus causing a shift in the reading frame of the gene. This shift leads to a change in the reading frame of all parts of a gene that are downstream from the mutation leading to a premature stop codon, and thus to a truncated protein product.
- Functional genomics** the development and implementation of technologies to characterize the mechanisms through which genes and their products function and interact with each other and with the environment.
- Gain-of-function mutation** a mutation that produces a protein that takes on a new or enhanced function.
- Gene** the fundamental unit of heredity; in molecular terms, a gene comprises a length of DNA that encodes a functional product, which may be a polypeptide (a whole or constituent part of a protein or an enzyme) or a ribonucleic acid. It includes regions that precede and follow the coding region as well as introns and exons. The exact boundaries of a gene are often ill defined since many promoter and enhancer regions dispersed over many kilobases may influence transcription.
- Gene-based therapy** refers to all treatment regimens that employ or target genetic material. This includes (1) *transfection* (introducing cells whose genetic make-up is modified), (2) *antisense* therapy, and (3) *naked DNA* vaccination.
- Gene editing** refers to correction of mutation in an endogenous gene, rather than silencing or replacing it; in most cases it is targeted at highly homologous recombination or patch-specific defects in a gene. A number of efficient tools are in use—zinc finger nucleases, TALENs, and most noteworthy CRISPR/Cas9 techniques.
- Gene expression** the process through which a gene is activated at a particular time and place so that its functional product is produced—transcription into mRNA followed by translation into protein.
- Gene expression profile** the pattern of changes in the expression of a specific set of genes that is relevant to a disease or treatment. The detection of this pattern depends upon the use of specific gene expression measurement technique.
- Gene family** a group of closely related genes that make similar protein products, for example, RAS-MAPK family of multiple genes.
- Gene regulatory network** a functional map of the relationships between a number of different genes and gene products (proteins), regulatory molecules, etc. that define the regulatory response of a cell with respect to a particular physiological function.
- Gene therapy** a therapeutic medical procedure that involves either replacing/manipulating or supplementing nonfunctional genes with healthy genes. Gene therapy can be targeted to somatic (body) or germ (egg and sperm) cells. In *somatic gene therapy* the recipient's genome is changed, but the change is not passed along to the next generation. In *germ-line gene therapy*, the parent's egg or sperm cells are changed with the goal of passing on the changes to their offspring.
- Genetics** refers to the study of heredity, gene, and genetic material. In contrast to genomics, the genetics is traditionally related to lower throughput, smaller scale emphasis on single genes, rather than on studying structure, organization, and function of many genes.
- Genetic architecture** refers to the full range of genetic effects on a trait. Genetic architecture is a moving target that changes according to gene and genotype frequencies, distributions of environmental factors, and such biological properties as age and sex.
- Genetic code** the relationship between the order of nucleotide bases in the coding region of a gene and the order of amino acids in the polypeptide product. It is universal, triplet, nonoverlapping code such that each set of three bases (termed a codon) specifies which of the 20 amino acids is present in the polypeptide chain product of a particular position.



- Genetic counseling** an important process for individuals and families who have a genetic disease or who are at risk for such a disease. Genetic counseling provides patients and other family members information about their condition and helps them make informed decisions.
- Genetic determinism** the unsubstantiated theory that genetic factors determine a person's health, behavior, intelligence, or other complex attributes.
- Genetic discrimination** unfavorable discrimination of an individual, a family, community, or an ethnic group on the basis of genetic information. Discrimination may include societal segregation, political persecution, opportunities for education and training, lack or restricted employment prospects, and adequate personal financial planning, for example, life insurance and mortgage.
- Genetic engineering** the use of molecular biology techniques, such as restriction enzymes, ligation, and cloning, to transfer genes among organisms (also known as *recombinant DNA cloning*).
- Genetic epidemiology** a field of research in which correlations are sought between phenotypic trends and genetic variation across population groups.
- Genetic map** a map showing the positions of genetic markers along the length of a chromosome relative to each other (genetic map) or in absolute distances from each other.
- Genetic screening** testing a population group to identify a subset of individuals at high risk for having or transmitting a specific genetic disorder.
- Genetic susceptibility** predisposition to a particular disease due to the presence of a specific allele or combination of alleles in an individual's genome.
- Genetic test** an analysis performed on human DNA, RNA, genes, and/or chromosomes to detect heritable or acquired genotypes. *Genetic testing* strictly refers to testing for a specific chromosomal abnormality or a DNA (nuclear or mitochondrial) mutation already known to exist in a family member. This includes diagnostic testing (postnatal or prenatal), presymptomatic, or predictive genetic testing or for establishing the carrier status. The individual concerned should have been offered full information on all aspects of the genetic test through the process of *nonjudgmental and nondirective* genetic counseling. Most laboratories require a formal fully informed signed consent before carrying out the test. Genetic testing commonly involves DNA/RNA-based tests for single gene variants, complex genotypes, acquired mutations, and measures of gene expression. Epidemiologic studies are needed to establish clinical validity of each method to establish sensitivity, specificity, and predictive value.
- Genome** the complete set of chromosomal and extra-chromosomal DNA/RNA of an organism, a cell, an organelle, or a virus.
- Genome** the complete set of genetic (hereditary) material in an organism; genomics refers to the study of entire genetic constitution and function in an organism.
- Genomics** the study of the genome and its action. The term is commonly used to refer to large-scale, high-throughput molecular analyses of multiple genes, gene products, or regions of genetic material (DNA and RNA). The term also includes the comparative aspect of genomes of various species, their evolution, and how they relate to each other (see *comparative genomics*).
- Genomic drugs** drugs based on molecular targets; genomic knowledge of the genes involved in diseases, disease pathways, and drug response.
- Genomic imprinting** an epigenetic phenomenon in which expression of certain genes depends on the parent of origin. Imprinted genes are silenced so that respective genes are expressed only in the presence of nonimprinted allele inherited from the mother or father.
- Genomic instability** an increased tendency of the GENOME to acquire MUTATIONS when various processes involved in maintaining and replicating the genome are dysfunctional.
- Genomic profiling** complete genomic sequence of an individual including the expression profile. This would be targeted to specific requirements, for example, most common complex diseases (diabetes, hypertension, and coronary heart disease).
- Genotype** the genetic constitution of an organism; commonly used in reference to a specific disease or trait.
- Germ-line mosaic (germinal mosaic, gonadal mosaic, gonosomal mosaic)** an individual who has a subset of germ-line cells carrying a mutation which is not found in other germ-line cells.
- Germ-line mutation** a gene change in the body's reproductive cells (egg or sperm) which becomes incorporated into the DNA of every cell in the body of offspring; germ-line mutations are passed on from parents to offspring, also called *hereditary mutation*.
- Green fluorescent protein (GFP)** a protein that exhibits bright green fluorescence when exposed to light in the blue or ultraviolet range, frequently used as a reporter in cell and molecular biology.
- GWAS** refers to the *genome-wide association study*, one of the most commonly used research methods for selecting genomic variants (single nucleotide polymorphisms, SNPs) in linkage disequilibrium with a complex/multifactorial condition. Despite its limitations, the outcome of GWAS is frequently misquoted and misinterpreted for clinical use. In most cases, the GWAS results simply indicate statistical data for genetic association and do not in any way imply pathogenic importance of any or group of genomic variants. However, the GWAS data contribute to assessing the missing heritability for a complex trait.
- Haploid** describing a cell (typically a gamete) with only a single copy of each chromosome (i.e., 23 in man).
- Haploinsufficiency** this occurs when an organism has only one functional copy of the gene, while the other copy is inactivated by a point mutation or pathogenic deletion or duplication. The functional copy alone is unable to produce the wild type, leading to the disease state.
- Haplotype** a series of closely linked loci on a particular chromosome inherited together as a block.
- Hemodynamics** the forces involved in circulating blood through the cardiovascular system. The heart adapts its hemodynamic performance to the needs of the body, increasing its output of blood when muscles are working and decreasing output when the body is at rest.
- Heterochromatin** nonexpressed eukaryotic chromatin that is highly condensed.
- Heteroplasmy** the presence of a mixture of more than one type of mitochondrial genome within a cell, usually wild-type and mutated mtDNA.
- Heterozygosity** the presence of different alleles of a gene in one individual or in a population—a measure of genetic diversity.
- Heterozygote** refers to a particular allele of a gene at a defined chromosome locus. A heterozygote has a different allelic form of the gene in each of the two homologous chromosomes.



- Holter monitoring** a technique for the continuous recording of electrocardiographic (ECG) signals, usually over 24 hours, to detect and diagnose ECG changes (also called ambulatory monitoring).
- Homology** similarity between two sequences due to their evolution from a common ancestor, often referred to as *homologs*.
- Homoplasmy** a cell in which all copies of the mtDNA are identical. The homoplasmic copies may be normal or mutated.
- Homozygote** refers to same allelic form of a gene on each of the two homologous chromosomes.
- Hormone** a molecular that mediates intercellular signaling, usually at a distance; it is secreted by one cell type and induces a physiological response in another cell, often at a distance.
- Human gene transfer** the process of transferring genetic material (DNA or RNA) into a person; an experimental therapeutic procedure to treat certain health problems by compensating for defective genes, producing a potentially therapeutic substance, or triggering the immune system to fight disease. This may help improve genetic disorders, particularly those conditions that result from inborn errors in a single gene (e.g., sickle cell anemia, hemophilia, and cystic fibrosis), and with complex disorders, such as cancer and heart disease, and certain infectious diseases, such as HIV/AIDS.
- Human Genome Project** a program to determine the sequence of the entire 3 billion bases of the human genome.
- Identity by descent (IBD)** alleles in an individual or in two people who are identical because they have been inherited from the same common ancestor, as opposed to *identity by state (IBS)*, which is coincidental possession of similar alleles in unrelated individuals (see *Consanguinity*).
- Immunogenomics** refers to the study of organization, function, and evolution of vertebrate defense genes, particularly those encoded by the major histocompatibility complex (MHC) and the leukocyte receptor complex (LRC). Both complexes form integral parts of the immune system. The MHC is the most important genetic region in relation to infection and common disease, such as autoimmunity. Driven by pathogen variability, immune genes have become the most polymorphic loci known, with some genes having over 500 alleles. The main function of these genes is to provide protection against pathogens and they achieve this through complex pathways for antigen processing and presentation.
- Immunoglobulin** commonly called an antibody, produced by the immune system that specifically binds to a foreign agent, an antigen. Typically, an antibody has a Y-shaped structure and neutralizes the pathogen or makes it ineffective through other components of the immune system.
- Imprinting** see genome imprinting.
- Indel** short insertions or deletions of nucleotides in the DNA.
- Induced pluripotent stem cells** a type of pluripotent stem cell that originates from differentiated normal somatic cells, propagating indefinitely or differentiated into other cell types; current application of individualized pluripotent stem cells is promising for treating selected genetic diseases.
- Informatics** the study of the application of computer and statistical techniques to the management of information. In genome projects, informatics includes the development of methods to search databases quickly, to analyze DNA sequence information, and to predict protein sequence and structure from DNA sequence data. In clinical practice, increasingly importance of *clinical informatics* is recognized.
- Innate immune response** the first line of defense of mammals against invading pathogens.
- In silico** a computational method for establishing the biological significance through computational simulation; the term was developed in relation to *in vivo* (internal) or *in vitro* (external).
- In situ hybridization** hybridization of a labeled nucleic acid to a target nucleic acid which is typically immobilized on a microscopic slide, such as DNA of denatured metaphase chromosomes [as in *fluorescent in situ hybridization (FISH)*] or the RNA in a section of tissue [as in *tissue in situ hybridization (TISH)*].
- Intron** a noncoding sequence within eukaryotic genes which separates the exons (coding regions). Introns are spliced out of the messenger RNA molecule created from a gene after transcription, prior to protein translation (protein synthesis).
- In vitro** (Latin) literally “in glass,” meaning outside of the organism in the laboratory, for example, a tissue culture.
- In vivo** (Latin) literally “in life,” meaning within a living organism.
- Ischemia** insufficient blood flow to tissue due to blockage in the blood flow through the arteries, commonly refers to ischemic heart disease (IHD).
- Isoforms/Isozymes** alternative forms of a protein/enzyme, for example, protease inhibitor forms of the alpha-1 antitrypsin—ZZ, MZ, SZ, MM, or SS.
- Karyotype** refers to the set of chromosomes in a cell of an individual; described as per internationally agreed cytogenetic symbols and nomenclature for numerical and structural changes.
- Kinase** an enzyme that transfers a phosphoryl group from high-energy donor molecules; for example, ATP to specific substrates. The protein kinases transfer the phosphate to a target protein.
- Knock-out** a technique used primarily in mouse genetics to inactivate a particular gene in order to define its function.
- Linkage** the phenomenon whereby pairs of genes which are located in close proximity on the same chromosome tend to be coinherited; *linkage analysis*: a process of locating genes on the chromosome by measuring recombination rates between phenotypic and genetic markers (see *LOD score*).
- Linkage disequilibrium** the nonrandom association in a population of alleles at nearby loci, commonly used in genetic or genomic association analysis for a particular complex disease, for example, HLAB27 in association with ankylosing spondylitis. See also GWAS.
- Locus** the specific site on a chromosome at which a particular gene or other DNA landmark is located.
- LOD score** a measure of likelihood of genetic linkage between loci; a LOD score greater than +3 is often taken as evidence of linkage; one that is less than -2 often taken as evidence against linkage.
- Loss-of-function mutation** a mutation that decreases the production or function (or both) of the gene product.
- Loss of heterozygosity (LOH)** loss of alleles on one chromosome detected by assaying for markers for which an individual is constitutionally heterozygous.
- Lyonization** introduced by Mar Lyon, refers to the process of random X chromosome inactivation in the early embryonic stage of developing mammals.

- Marker** a specific feature at an identified physical location on a chromosome, whose inheritance can be followed. The position of a gene implicated in a particular phenotypic effect can be defined through its linkage to such markers.
- Meiosis** reductive cell division occurring exclusively in testis and ovary and resulting in the production of haploid cells, including sperm cells and egg cells.
- Mendelian genetics** classical genetics, which focuses on *monogenic* genes with high *penetrance*. The Mendelian genetics is a true *paradigm* and is used in discussing the mode of inheritance (see *Monogenic disease*).
- Messenger RNA (mRNA)** RNA molecules that are synthesized from a DNA template in the nucleus (a gene) and transported to ribosomes in the cytoplasm where they serve as a template for the synthesis of protein (translation).
- Micro RNAs** a short endogenous RNA of ~21 nucleotides that regulates gene expression through RNA interference.
- Microsatellite DNA** small array (often less than 0.1 kb) of short tandemly repeated DNA sequences.
- Minisatellite DNA** an intermediate size array (often 0.1–20 kb long) of short tandemly repeated DNA sequences. *Hypervariable minisatellite* DNA is the basis of DNA fingerprinting and many VNTR markers.
- Missense mutation** substitution of a single DNA base that results in a codon that specifies an alternative amino acid.
- Mitochondria** cellular organelles present in eukaryotic organisms which enable aerobic respiration and generate the energy to drive cellular processes. Each mitochondria contains a small amount of circular DNA encoding a small number of genes (approximately 50); the term *mitochondriopathies* refers to the group of mitochondrial diseases caused by either mutations in mtDNA or nDNA genes.
- Mitosis** cell division in somatic cells.
- Modifier gene** a gene whose expression can influence a phenotype resulting from mutation at another locus.
- Molecular genetic screening** screening a section of the population known to be at a higher risk to be heterozygous for one of the mutations in the gene for a common autosomal recessive disease, for example, screening for cystic fibrosis in the North European populations and beta-thalassemia in the Mediterranean and Middle-East population groups.
- Molecular genetic testing** molecular genetic testing for use in patient diagnosis, management, and genetic counseling; this is increasingly used in presymptomatic (predictive) genetic testing of “at-risk” family members using a previously known disease-causing mutation in the family.
- Mosaic** a genetic mosaic is an individual who has two or more genetically different cell lines derived from a single zygote.
- Multifactorial disease** any disease or disorder caused by interaction of multiple genetic (polygenic) and environmental factors. See also *GWAS*.
- Mutation** a heritable alteration in the DNA sequence.
- Nanobody** a recombinant, single domain antibody derived from a camelid heavy-chain antibody that consists only of a dedicated antigen-binding V<sub>H</sub>H domain.
- Natural selection** the process whereby some of the inherited genetic variations within a population will affect the ability of individuals to survive to reproduce (*fitness*).
- Neutral mutation** a change or alteration in DNA sequence which has no phenotypic effect (or has no effect on fitness).
- Newborn screening** performed in newborns in state public health programs to detect certain genetic diseases for which early diagnosis and treatment are available.
- Noncoding sequence** a region of DNA that is not translated into protein. Some noncoding sequences are regulatory portions of genes, others may serve structural purposes (telomeres and centromeres), while others may not have any function.
- Nonsense mutation** substitution of a single DNA base that leads in a stop codon, thus leading to the truncation of a protein.
- Northern blot hybridization** a form of molecular hybridization in which target consists of RNA molecules that have been size fractioned by gel electrophoresis and subsequently transferred to a membrane.
- Nucleotide** a subunit of the DNA or RNA molecule. A nucleotide is a base molecule (adenine, cytosine, guanine, and thymine in the case of DNA), linked to a sugar molecule (deoxyribose or ribose) and phosphate groups.
- Oncogene** an acquired mutant form of a gene which acts to transform a normal cell into a cancerous one.
- OMIM** acronym for McKusick’s Online Mendelian Inheritance in Man, a regularly updated electronic catalog of inherited human disorders and phenotypic traits accessible on NCBI network. Each entry is designated by a number (*MIM number*).
- PEGylation** the process of covalently attaching polyethylene glycol (PEG) to another molecule, such as a protein or an aptamer.
- Penetrance** the likelihood that a person carrying a particular mutant gene will have an altered phenotype (see *phenotype*).
- Personalized medicine** refers to individualized tailoring of treatment based on biomarkers or any other characteristics; this is used in conjunction with precision medicine based on genetic or genomic factors, such as genomic polymorphic variants.
- PFGE (pulse field gel electrophoresis)** a form of gel electrophoresis which permits size fractionation of large DNA molecules, not used now except in exceptional situation.
- Pharmacodynamics** refers to investigations into the mechanism of drug action in relation to drug concentration and effect; in contrast *pharmacokinetics* refers to the rate of drug absorption and distribution to different organs.
- Pharmacogenetics (syn: pharmacogenomics)** the identification of the genes which influence individual variation in the efficacy or toxicity of therapeutic agents, and the application of this information in clinical practice.
- Phenotype** the clinical and/or any other manifestation or expression, such as a biochemical immunological alteration, of a specific gene or genes, environmental factors, or both.
- Pleiotropy** refers to genetic variation in relation to seemingly unrelated phenotypic traits.
- Pluripotency** the ability of a cell to differentiate into any cell of the three germ layers—ectoderm, mesoderm, and endoderm.
- Point mutation** the substitution of a single DNA base in the normal DNA sequence.
- Polygenic trait or character** a character or trait determined by the combined action of a number of loci, each with a small effect.
- Polymerase chain reaction (PCR)** a molecular biology technique developed in the mid-1980s through which specific DNA segments may be amplified selectively.
- Polymorphism** the stable existence of two or more variant allelic forms of a gene within a particular population, or among different populations.

- Positional cloning** the technique through which candidate genes are located in the genome through their coinheritance with linked markers.
- Posttranscriptional modification** a series of steps through which protein molecules are biochemically modified within a cell following synthesis by translation of messenger RNA. A protein may undergo a complex series of modifications in different cellular compartments before its final functional form is produced.
- Predictive testing** determines the probability that a healthy individual with or without a family history of a certain disease might develop that disease.
- Predisposition, genetic** increased susceptibility to a particular disease due to the presence of one or more gene mutations, and/or a combination of alleles (haplotype), not necessarily abnormal, that is associated with an increased risk for the disease, and/or a family history that indicates an increased risk for the disease; *predisposition testing* is a test for a genetic predisposition (incompletely *penetrant* conditions). Not all people with a positive test result will manifest the disease during their lifetimes.
- Preimplantation genetic diagnosis (PIGD)** used following in vitro fertilization to diagnose a genetic disease or condition in a preimplantation embryo.
- Prenatal diagnosis** used to diagnose a genetic disease or condition in a developing fetus.
- Presymptomatic test** predictive testing of individuals with a family history. Historically, the term has been used when testing for diseases or conditions, such as Huntington's disease, where the likelihood of developing the condition (known as *penetrance*) is very high in people with a positive test result.
- Primer** a short nucleic acid sequence, often a synthetic oligonucleotide, which binds specifically to a single strand of a target nucleic acid sequence and initiates synthesis, using a suitable polymerase, of a complementary strand; synthetic primers are commonly used in PCR-based molecular analysis.
- Prion protein** a type of disease-causing protein, for example, scrapie and frontotemporal dementia.
- Probe** a DNA or RNA fragment which has been labeled in some way and used in a *molecular hybridization* assay to identify closely related DNA or RNA sequences.
- Prokaryote** an organism or cell lacking a nucleus and other membrane-bounded organelles. Bacteria are prokaryotic organisms.
- Promoter** a combination of short sequence elements to which RNA polymerase binds in order to initiate transcription of a gene.
- Protein** a protein is the biological effector molecule encoded by sequences of a gene. A protein molecule consists of one or more polypeptide chains of amino acid subunits. The functional action of a protein depends on its three-dimensional structure, which is determined by its amino acid composition.
- Proteome** all of the proteins present in a cell or organism; *proteomics*—the development and application of techniques to investigate the protein products of the genome and how they interact to determine biological functions.
- Proto-oncogene** a cellular gene which when mutated is inappropriately expressed and becomes an oncogene.
- Pseudoautosomal region (PAR)** a region on the tips of mammalian X chromosomes which is involved in recombination during male meiosis.
- Pseudogene** a DNA sequence which shows a high degree of sequence homology to a nonallelic functional gene but which is itself nonfunctional.
- Recessive** an allele that has no phenotypic effect in the heterozygous state.
- Recombinant DNA technology** the use of molecular biology techniques, such as restriction enzymes, ligation, and cloning to transfer genes, among organisms (see genetic engineering).
- Regulatory mutation** a mutation in a region of the genome that does not encode a protein but affects the expression of a gene; *regulatory sequence*—a DNA sequence to which specific proteins bind to activate or repress the expression of a gene.
- Repeat sequences** a stretch of DNA bases that occurs in the genome in multiple identical or closely related copies.
- Replication** a process by which a new DNA strand is synthesized by copying an existing strand, using it as a template for the addition of a complementary bases, catalyzed by a DNA polymerase enzyme.
- Reproductive cloning** techniques aimed at the generation of an organism with an identical genome to an existing organism.
- Restriction enzymes** a family of enzymes derived from bacterial that cut DNA at specific sequences of bases.
- Restriction fragment length polymorphism (RFLP)** a polymorphism due to difference in size of allelic restriction fragments as a result of restriction site polymorphism.
- Ribonucleic acid (RNA)** a single stranded nucleic acid molecule comprising a linear chain made up from four nucleotide subunits (A, C, G, and U). There are three types of RNA: messenger, transfer, and ribosomal.
- Risk communication** an important aspect of genetic counseling which involves pedigree analysis, interpretation of the inheritance pattern, genetic risk assessment, and explanation to the family member (or the family).
- RNA interference** a mechanism of posttranscriptional gene regulation; short double-stranded RNA molecules trigger the degradation of a homologous RNA by the RNA-induced silencing complex (RISC).
- RT-PCR (reverse transcriptase-PCR)** a PCR reaction in which the target DNA is a cDNA copied by reverse transcriptase from an mRNA source.
- Screening** carrying out of a test or tests, examination(s), or procedure(s) in order to expose undetected abnormalities, unrecognized (incipient) diseases, or defects: examples are early diagnosis of cancer using mass X-ray mammography for breast cancer and cervical smears for cancer of the cervix.
- Second messenger** refers to signaling molecules that transfer extracellular events, such as binding of a hormone to a receptor on the cell surface to a target molecule inside the cell, for example, cyclic AMP.
- Segregation** the separation of chromosomes (and the alleles they carry) during meiosis; alleles on different chromosomes segregate randomly among the gametes (and the progeny).
- Sensitivity (of a screening test)** extent (usually expressed as a percentage) to which a method gives results that are free from false negatives; the fewer the false negatives, the greater the sensitivity. Quantitatively, sensitivity is the proportion of truly diseased persons in the screened population who are identified as diseased by the screening test.
- Sex chromosome** the pair of chromosomes that determines the sex (gender) of an organism. In man, one X and one Y chromosomes constitute a male compared to two X chromosomes in a female.

- Sex selection** preferential selection of the unborn child on the basis of the gender for social and cultural purposes. However, this may be acceptable for medical reasons, for example, to prevent the birth of a male assessed to be at risk for an X-linked recessive disease. For further information visit [http://www.bioethics.gov/topics/sex\\_index.html](http://www.bioethics.gov/topics/sex_index.html).
- Short hairpin RNA (shRNA)** a self-complementary double-stranded RNA molecule that folds back on itself in a tight hairpin turn; sh RNAs are generated after the intracellular expression of triggers for RNAi-mediated silencing. The cellular protein Dicer processes the shRNA into small interfering RNAs (siRNAs).
- Shotgun sequencing** a cloning method in which total genomic DNA is randomly sheared and the fragments ligated into a cloning vector, also referred to as “shotgun” cloning.
- Signal transduction** the molecular pathways through which a cell senses changes in its external environment and changes its gene expression patterns in response.
- Silent mutation** substitution of a single DNA base that produces no change in the amino acid sequence of the encoded protein.
- Single-nucleotide polymorphism (SNP)** a common variant in the genome sequence; the human genome contains about 10 million SNPs.
- Small interfering RNA (siRNA)** a short double-stranded RNA molecule that can be used to silence the gene expression by RNA interference.
- Somatic** all of the cells in the body which are not gametes (germ-line).
- Southern blot hybridization** a form of molecular hybridization in which the target nucleic acid consists of DNA molecules that have been size fractionated by gel electrophoresis and subsequently transferred to a nitrocellulose or nylon membrane.
- Splicing** a process by which introns are removed from a messenger RNA prior to translation and the exons adjoined.
- Stem cell** a cell which has the potential to differentiate into a variety of different cell types depending on the environmental stimuli it receives.
- Stop codon** a codon that leads to the termination of a protein rather than to the addition of an amino acid. The three stop codons are TGA, TAA, and TAG.
- Syncope** fainting, loss of consciousness, or dizziness which may be due to a transient disturbance of cardiac rhythm (arrhythmia) or other causes.
- Synteny** a large group of genes that appear in the same order on the chromosomes of two different species.
- Systems biology** refers to simultaneous measurement of thousands of molecular components (such as transcripts, proteins, and metabolites) and integrates these disparate data sets with clinical end points, in a biologically relevant manner; this model can be applied in understanding the etiology of disease.
- Telomere** the natural end of the chromosome.
- Therapeutic cloning** the generation and manipulation of stem cells with the objective of deriving cells of a particular organ or tissue to treat a disease.
- Transcription** the process through which a gene is expressed to generate a complementary RNA molecule on a DNA template using RNA polymerase.
- Transcription factor** a protein which binds DNA at specific sequences and regulates the transcription of specific genes.
- Transcriptome** the total messenger RNA expressed in a cell or tissue at a given point in time.
- Transfection** a process by which new DNA is inserted in a eukaryotic cell allowing stable integration into the cell's genome.
- Transformation** introduction of foreign DNA into a cell and expression of genes from the introduced DNA; this does not necessarily include integration into host cell genome.
- Transgene** a gene from one source that has been incorporated into the genome of another organism.
- Transgenic animal/plant** a fertile animal or plant that carries an introduced gene(s) in its germ-line.
- Translation** a process through which a polypeptide chain of amino acid molecules is generated as directed by the sequence of a particular messenger RNA sequence.
- Tumor suppressor gene** a gene which serves to protect cells from entering a cancerous state; according to Knudson's “two-hit” hypothesis, both alleles of a particular tumor suppressor gene must acquire a mutation before the cell will enter a transformed cancerous state.
- Uniparental disomy** occurs when a person receives two copies of a chromosome from one parent and no copy from the second parent; an important phenomenon in genetic imprinting where recognizable phenotypes result in case of imprinted genes.
- Ventricular fibrillation (VF)** very fast, chaotic, quivering heart contractions that start in the ventricles. During VF, the heart does not beat properly. This often results in fainting. If left untreated, it may result in cardiac arrest. Blood is not pumped from the heart to the rest of the body. Death will occur if defibrillation is not initiated within 6 minutes from the onset of VF.
- Ventricular tachycardia (VT)** a rapid heart rate that starts in the ventricles. During VT, the heart does not have time to fill with enough blood between heart beats to supply the entire body with sufficient blood. It may cause dizziness and light headedness.
- Virulence factors** refer to molecules, for example, proteins that enable a microorganism to establish itself within a host and enhance its potential to cause disease.
- Western blotting** a process in which proteins are size-fractionated in a polyacrylamide gel prior to transfer to a nitrocellulose membrane for probing with an antibody.
- X-chromosome inactivation** random inactivation of one of the two X chromosomes in mammals by a specialized form of genetic imprinting (see *Lyonization*); most likely explanation of an affected female with an X-linked recessive gene mutation where the normal X was randomly inactivated.
- Yeast artificial chromosome (YAC)** an artificial chromosome produced by combining large fragments of foreign DNA with small sequence elements necessary for chromosome function in yeast cells.
- Zinc finger** a polypeptide motif stabilized by binding a zinc atom and confers on proteins an ability to bind specifically to DNA sequences; commonly found in transcription factors.
- Zymogen** an inactive precursor of a protein (enzyme) that must undergo limited proteolysis to become fully active, also known as proprotein or proenzyme.





# Index

Note: Page numbers followed by “f,” “t,” and “b” refer to figures, tables, and boxes respectively.

## A

- Abacavir, 525  
  in HIV infection, 527  
Abelson oncogene (*ABL*), 428–429  
  *ABL1* gene, 497  
ACC/AHA. *See* American College of  
  Cardiology/American Heart Association  
  (ACC/AHA)  
“Acceptable mismatch” strategy, 321–322  
Aceruloplasminaemia, 482  
Acetyl-CoA, 43  
Achondroplasia, 100  
ACMG. *See* American College of Medical  
  Genetics and Genomics (ACMG)  
ACPA. *See* Anticitrullinated peptide  
  antibodies (ACPA)  
ACTH. *See* Adrenocorticotrophic hormone  
  (ACTH)  
Actin microfilaments, 21  
Activity-regulated cytoskeleton-associated  
  protein (ARC), 492  
Acute brain injury, 290  
Acute encephalitis, 412  
Acute HCV infection, 407  
Acute inflammation, 389  
  groups of serum inflammatory markers,  
    390f  
  molecular pathology, 390–391  
Acute kidney injury (AKI), 437–438. *See also*  
  Chronic kidney disease (CKD)  
  candidate gene association studies, 438  
  GWASs, 438  
Acute lymphatic leukemia (ALL), 427, 429  
Acute myeloid leukemia (AML), 427  
Acute necrotizing encephalopathy (ANE),  
  412  
Acute Q fever, 414–415  
Acute stroke therapy, 189  
AD. *See* Alzheimer’s disease (AD)  
ADA. *See* Adenosine deaminase (ADA);  
  American Diabetes Association (ADA)  
ADAM33, 359  
Adapter proteins, 19  
Adaptive immune system, 379  
Adaptive immunity, 379, 407  
Adenine (A), 4  
Adenomatous polyposis coli protein (APC  
  protein), 21  
Adenosine 5′-monophosphate (AMP), 19  
Adenosine deaminase (ADA), 99  
Adenosine triphosphate (ATP), 19, 475  
Adhesion junctions, 21  
Adiposity, 78  
Adjuvant therapy, 526  
ADPKD. *See* Autosomal dominant polycystic  
  kidney disease (ADPKD)  
ADRB2, 359  
Adrenaline, 148  
Adrenocorticotrophic hormone (ACTH),  
  80–81, 217  
Adult respiratory distress syndrome (ARDS),  
  390  
Adverse antiepileptic drug reactions,  
  299–300  
*Aedes aegypti*, 410  
Afibrinogenemia, 348  
Age-related macular degeneration (AMD),  
  514–515  
Aging, 392  
Agouti-related protein (AGRP), 77  
aHUS. *See* Atypical hemolytic uremic  
  syndrome (aHUS)  
Airway epithelial homeostasis, 362–363  
AIS. *See* Androgen insensitivity syndrome  
  (AIS)  
AJ. *See* Ashkenazi Jewish (AJ)  
AKAP9, 139  
AKI. *See* Acute kidney injury (AKI)  
ALAS-2. *See*  $\delta$ -aminolevulinic acid synthase-2  
  (ALAS-2)  
Albright’s hereditary osteodystrophy.  
  *See* Type 1A PHP  
ALCHEMIST trial, 531–532  
Aldosterone-producing adenoma (APA), 230  
Aldosterone-to-renin-activity-ratio (ARR),  
  230  
Alectinib, 498  
Alignment, 65  
ALK. *See* Anaplastic lymphoma kinase  
  (ALK)  
ALL. *See* Acute lymphatic leukemia (ALL)  
Allele-specific HLA expression, 316–317  
Allelic data, 68  
Alloantigen recognition, 319  
Allogenomics mismatch score (AMR), 442  
Allorecognition in HLA, 317–318  
ALOX5AP. *See* Arachidonate 5-lipoxygenase-  
  activating protein gene (*ALOX5AP*)  
21- $\alpha$  hydroxylase, 233  
5- $\alpha$  reductase, 242  
 $\alpha$  thalassemia, 332–334. *See also*  $\beta$   
  thalassemia  
  classification, 332  
  clinical phenotypes, 332  
laboratory diagnosis  
  of  $\alpha$ -deletions, 332–333  
  of point mutations, 332–333  
  of triplications, 332–333  
molecular pathology  
  of deletional, 332  
  of nondeletional, 332  
peripheral blood in patient with  
  homozygous Hb Sallanches, 333f  
TI, 333–334  
 $\alpha$ -globin genes, 327, 334. *See also*  $\beta$ -globin  
  gene  
   $\alpha$  sequencing, 333  
  cluster, 327  
 $\alpha$ -globin variants, 335–336  
 $\alpha$ -helical segments (A–H), 327  
17 $\alpha$ -hydroxylase deficiency, 178–180  
  enzyme effect affecting steroid hormone  
  synthesis, 180f  
  genetics, 180  
  pathophysiology, 180  
 $\alpha$ -melanocyte-stimulating hormone  
  ( $\alpha$ -MSH), 77  
Alpha-synuclein gene, 461–462  
Alport syndrome (AS), 439  
ALS. *See* Amyotrophic lateral sclerosis (ALS)  
Altered repertoire mechanism, 313  
“Altered self” antigens, 312–313  
Alzheimer’s disease (AD), 63, 447, 449–450  
Ambulatory electrocardiography, 156–157  
AMD. *See* Age-related macular degeneration  
  (AMD)  
American College of Cardiology/American  
  Heart Association (ACC/AHA), 249  
American College of Medical Genetics and  
  Genomics (ACMG), 61–62, 68  
American Diabetes Association  
  (ADA), 270  
Amino acid changes, 66–67  
Amino acyl-tRNA-synthases, 38–39  
AML. *See* Acute myeloid leukemia (AML)  
Amniocentesis, 331  
AMP. *See* Adenosine 5′-monophosphate  
  (AMP)  
AMP/ATP ratio, 32  
Amplification methods, 403  
Amplification refractory mutation system  
  (ARMS), 331  
AMR. *See* Allogenomics mismatch score  
  (AMR); Antibody-mediated rejection  
  (AMR)

- Amyloid beta precursor protein gene (*APP*), 190
- Amyloid oligomers, 454
- Amyloidosis, 393–394
- Amyotrophic lateral sclerosis (ALS), 447
- Anabolic signaling, 391
- Anaphylactic reaction, 342
- Anaplastic lymphoma kinase (ALK), 498  
inhibitors, 523
- Andersen–Tawil syndrome, 138, 141
- Androgen insensitivity syndrome (AIS), 112, 242
- Androgen overproduction, 114–115
- Androgen receptor (AR), 112–113
- Androgen resistance/insensitivity syndrome, 240–242
- Androgen response, disorders of, 112–113
- Androgen synthesis, disorders of, 111–112  
associated with adrenal dysfunction, 111  
associated without adrenal dysfunction, 112  
steroid hormone synthesis pathway with  
biochemical abnormalities, 112*f*
- ANE. *See* Acute necrotizing encephalopathy (ANE)
- Angiotensin II type 1 receptor, 320–321
- Angiotensin receptor neprilysin inhibitor (ARNI), 164
- Angiozyme, 514
- ANGPLT3 (Evinacumab), 260
- Ankylosing spondylitis (AS), 380  
and related spondyloarthropathies, 312
- Annexin A3 (*ANXA3*), 457
- Anomalous rectifying, 135
- Anorexigenic effect, 81
- Anosmin 1, 221
- Anti-mRNA strategies, 512–513
- Anti-Müllerian hormone (AMH) formation, 110
- Anti-PD-1 antibodies, 500
- Antiarrhythmic drugs, 165
- Antibody-dependent cell-mediated cytotoxicity, 319
- Antibody-mediated rejection (AMR), 318, 318*f*
- Anticholinergic medications, 471
- Anticitrullinated peptide antibodies (ACPA), 395
- Anticoagulation therapy, 344
- Antiepileptic  
drugs, 290, 299–300  
medication, 300
- Antigen recognition site (ARS), 315
- Antigen-presenting cell resistance (APCR), 344
- Antigen-presenting cells (APCs), 311–312, 406
- Anti–human leukocyte antigen antibodies, 320
- Antisense oligonucleotide (ASO), 99, 505, 512–514  
therapy, 513, 513*f*, 515*f*
- Antithrombin (AT), 341  
ATIII deficiency, 343
- ANXA3*. *See* Annexin A3 (*ANXA3*)
- AP2S1* gene, 225
- APA. *See* Aldosterone-producing adenoma (APA)
- APC protein. *See* Adenomatous polyposis coli protein (APC protein)
- APCR. *See* Antigen-presenting cell resistance (APCR)
- APCs. *See* Antigen-presenting cells (APCs)
- APECED. *See* Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED)
- apoB100. *See* Apolipoprotein B100 (apoB100)
- APOBEC3G. *See* Apolipoprotein B mRNA editing catalytic polypeptide 3G (APOBEC3G)
- APOE. *See* Apolipoprotein E gene (APOE)
- Apolipoprotein B mRNA editing catalytic polypeptide 3G (APOBEC3G), 405
- Apolipoprotein B100 (apoB100), 245–246
- Apolipoprotein E gene (APOE), 201
- Apolipoproteins, 245
- Apoptosis, 408
- APP*. *See* Amyloid beta precursor protein gene (*APP*)
- Apparent mineralocorticoid excess, 235–236
- Aptamers, 515–516
- Aquaporin 9 (*AQP9*), 457
- AR. *See* Androgen receptor (AR)
- Arachidonate 5-lipoxygenase-activating protein gene (*ALOX5AP*), 195, 197–198
- Arachidonic acid, 198
- ARC. *See* Activity-regulated cytoskeleton-associated protein (ARC)
- ARCN1* gene, 36
- ARDS. *See* Adult respiratory distress syndrome (ARDS)
- ARMS. *See* Amplification refractory mutation system (ARMS)
- ARNI. *See* Angiotensin receptor neprilysin inhibitor (ARNI)
- ARR. *See* Aldosterone-to-renin-activity-ratio (ARR)
- Array analysis, 83
- Array comparative genome hybridization, 123
- Arrhythmias, 137, 140–141, 156–157  
generation and propagation  
Brugada syndrome, 146–147  
CPVT, 145  
LQTS, 143–144  
in SQTS, 144  
increased automaticity, 137  
reentry, 137  
triggered activity, 137
- ARS. *See* Antigen recognition site (ARS)
- Arthritis, 394–395. *See also* Rheumatoid arthritis (RA)
- Artificial ribozyme, 514
- AS. *See* Alport syndrome (AS); Ankylosing spondylitis (AS)
- AS ONs. *See* Antisense oligonucleotide (ASO)
- ASDs. *See* Atrial septal defects (ASDs); Autism spectrum disorders (ASDs)
- Ashkenazi Jewish (AJ), 432–433
- ASO. *See* Antisense oligonucleotide (ASO)
- Aspergillus*, 416  
genetic variability of host and susceptibility, 416
- Aspergillus nidulans* complex, 416
- Asthma, 357. *See also* Bronchial asthma  
asthma-causing gene, 357  
asthma-related haplotypes, 362  
candidate genes, 362  
exacerbation, 353  
pathobiology, 360  
pathogenesis, 355  
pathophysiology, 363
- Astral microtubules, 41
- AT. *See* Antithrombin (AT); Ataxia-telangiectasia (AT)
- ATAC-seq. *See* Transposase-accessible chromatin sequencing (ATAC-seq)
- Ataxia, 472–480  
ataxia-telangiectasia, 479  
cerebellar degeneration, 476, 477*t*  
clinical and genetic description, 474*t*  
Freidreich's ataxia, 476  
genetics of, 475  
gene transcription and RNA, 475  
intranuclear inclusions, 475  
mitochondrial dysfunction, 475  
neuronal calcium homeostasis, 475  
transmembrane channel abnormalities, 475  
potential future targets for molecular therapy, 479
- SCA2, 476–478
- SCA3, 478
- SCA7, 478–479  
targeted molecular diagnosis and therapy  
diagnostic testing, 476
- Ataxia telangiectasia and rad3-related kinase (ATR kinase), 41
- Ataxia telangiectasia mutated kinase (ATM kinase), 41, 429
- Ataxia-telangiectasia (AT), 479  
therapy, 479
- Ataxin-3 gene (*ATXN3*), 478
- ATG. *See* Autophagy (ATG)
- ATG16L1. *See* Autophagy 16–like gene (*ATG16L1*)
- Atherosclerosis, 393–394
- Atherothrombotic cerebral infarction, 191–193
- ATM kinase. *See* Ataxia telangiectasia mutated kinase (ATM kinase)
- ATN-1 gene. *See* Atrophin-1 gene (ATN-1 gene)
- ATP. *See* Adenosine triphosphate (ATP)
- ATP13A2 protein, 483
- ATP7B protein, 36–37, 480–481
- ATR kinase. *See* Ataxia telangiectasia and rad3-related kinase (ATR kinase)
- Atrial septal defects (ASDs), 123
- Atrioventricular nodal reentry tachycardia (AVNRT), 146
- Atrophin-1 gene (ATN-1 gene), 483
- Attention deficit hyperactivity disorder (ADHD), 78

- ATXN2 gene, 478  
 ATXN3. *See* Ataxin-3 gene (ATXN3)  
 Atypical hemolytic uremic syndrome (aHUS), 438, 440  
 Atypical neuroaxonal dystrophy, 481  
 Autism spectrum disorders (ASDs), 489  
 Autoimmune endocrinopathy syndrome, 270  
 Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED), 415–416  
 Autoimmunity, 272–273  
 Autophagosomes, 38, 379  
 Autophagy (ATG), 38, 379  
   genes, 379  
 Autophagy 16–like gene (ATG16L1), 371–372  
 Autosomal dominant inheritance  
   LRRK2/PARK8, 462–463  
   SNCA/PARK1, 461–462  
   VPS35/PARK17, 463  
 Autosomal dominant polycystic kidney disease (ADPKD), 101–102  
 Autosomal recessive inheritance, 349  
   DJ1/PARK7, 463  
   GBA mutations, 463  
   genes implicated in Parkinson's disease, 464  
   Parkin/PARK2, 463  
   PINK1/PARK6, 463  
 Autosomal recessive spinal muscular atrophy, 29  
 Autosomal-dominant inheritance, 344  
 Autosomal-dominant Parkinson's disease, 463  
 Autosomal-dominant syndrome, 429  
 Autosomal-recessive adenomatous polyposis, 425  
 AVNRT. *See* Atrioventricular nodal reentry tachycardia (AVNRT)  
 Axoneme, 41–42
- B**  
 B-cell CLL, 427  
 B-type natriuretic peptide (BNP), 155  
 BAC/PAC clones, 91  
 Bardet–Biedl syndrome (BBS), 60, 82–83  
 Bariatric surgery, 84–85  
   in genetic obesity, 85  
 Base pairs (bps), 24–28  
 Basket trial design, 531, 531f  
 Basolateral domain, 20  
 BAV. *See* Bicuspid aortic valve (BAV)  
 Bayesian fine-mapping methodologies, 373  
 BBS. *See* Bardet–Biedl syndrome (BBS)  
 BCG vaccination, 355  
 BCG vaccines, 412  
 BCR. *See* Breakpoint cluster region (BCR)  
 Becker muscular dystrophy, 513  
 Beckwith–Wiedemann syndrome, 426  
 Benzodiazepines, 471  
 Bernard–Soulier syndrome, 349  
 $\beta$  thalassemia, 328–329. *See also*  $\alpha$  thalassemia  
   dominant, 330  
   gene deletions in, 330  
   laboratory diagnosis, 331  
   molecular basis of nondeletional, 329  
   molecular causes of, 330–331  
     silent and almost silent  $\beta$ Ts, 330  
     transacting mutations, 330  
     uniparental isodisomy/somatic deletion of  $\beta$ -globin gene, 331  
   molecular pathogenesis, 329  
   prevention, 331–332  
     noninvasive prenatal diagnosis, 332  
     PGD, 331–332  
     prenatal diagnosis, 331  
     screening for  $\beta$ -thalassemia trait, 331  
 $\beta$  thalassemia major ( $\beta$ TM), 328–329  
 $\beta$ -blockers, 134, 140–141  
 $\beta$ -catenin, 21  
 $\beta$ -globin gene, 330. *See also*  $\alpha$ -globin genes  
   cluster, 327  
   deletions restricted to, 330  
   expression and control, 328  
 $\beta$ -globin mRNA translation, mutants affecting, 329–330  
 $\beta$ -globin variants, 335–336  
 $\beta$ -hydroxylase deficiency, 178, 181  
 11 $\beta$ -hydroxylase deficiency (11 $\beta$ -OH deficiency), 114, 178–179  
   adrenal steroid synthesis, 179f  
   genetics, 178  
   pathophysiology, 179  
 3 $\beta$ -hydroxysteroid dehydrogenase deficiency (3 $\beta$ -HSD deficiency), 111, 114  
 11 $\beta$ -hydroxysteroid dehydrogenase gene (11 $\beta$ -HSD gene), 235  
 $\beta$ -promoter mutations, 328  
 Beta-propeller protein-associated neurodegeneration, 482  
 $\beta$ -thalassemia traits ( $\beta$ Ts), 328  
   screening for, 331  
   silent and almost silent, 330  
 Bevasiranib, 514–515  
 Bicuspid aortic valve (BAV), 123  
 Bile acid–sequestering agents, 259  
 Biochemical serum markers, 394  
 Biofluids, 457  
 Bioinformatics, 12, 65–66  
 Biomarkers, 47, 49–50, 298, 368–369, 392, 456–457  
 Biomolecules, 438  
 Bipolar affective disorder type I (BPI). *See* Manic depressive disorder  
 Bipolar disorder, 491–492  
 Bleeding disorder, 345–348. *See also*  
   Hemoglobin disorders (Hb disorders)  
   case study, 345  
   hemophilia A, 345–346  
   molecular basis of inherited  
     coagulopathies, 346–348  
     afibrinogenemia, 348  
     dysfibrinogenemia, 348  
     hemophilia B, 348  
     VWD, 346–347, 347f  
 Blepharophimosis, ptosis, and epicanthus inversus syndrome (BPES), 113  
 Blood. *See also* Coagulation; Hemoglobin (Hb)  
   glucose parameters, 275–276  
   plasma tests, 476, 478f  
   transfusion, 321  
 Blood pressure (BP), 169, 217  
   regulation, 169–170, 184  
 BMI. *See* Body mass index (BMI)  
 BNP. *See* B-type natriuretic peptide (BNP)  
 BOADICEA risk assessment, 427  
 Body fat, 73, 79  
 Body mass index (BMI), 73, 74f  
 Botulinum toxin therapy, 471  
 BP. *See* Blood pressure (BP)  
 BPES. *See* Blepharophimosis, ptosis, and epicanthus inversus syndrome (BPES)  
 BRAF, 497–498  
   inhibitors, 499  
   V600 mutations, 500  
   V600E, 497–498  
 Brain iron accumulation, neurodegeneration with, 481–483  
 BRCA1 gene, 424  
 BRCA2 gene, 424  
 Breakpoint cluster region (BCR), 428–429  
 Breast cancer, 73, 424–425  
 Brigatinib, 498  
 British “Serious Hazards of Transfusion” initiative, 322  
 Bronchial asthma, 353, 393–394  
   drugs in, 363  
   evolution of genomic research, 354f  
   genetic and genomic studies, 354–363  
   genomics, 353–354  
   management and treatment, 363  
     inflammation control, 363  
     symptomatic control, 363  
 Bronchodilators, 363  
 Brugada syndrome (BrS), 134–135, 146–147, 149  
   clinical risk assessment in, 149  
   generation and propagation of arrhythmia in, 146–147  
   targeted therapy in, 149  
 Bucket trial design. *See* Basket trial design  
 Bundles, 394  
 Burrows–Wheeler aligner package (bwa package), 65
- C**  
 C-peptide molecule of insulin, 268–269  
 C-reactive protein (CRP), 390–391  
 C-type natriuretic peptide acts (CNP acts), 103  
 C/EBP. *See* CCAAT enhancer-binding protein (C/EBP)  
 C3 (volanesorsen), 260  
 C4 genes. *See* Complement component 4 genes (C4 genes)  
 CA. *See* Cancer antigen (CA)  
 CA15–3 marker, 430  
 CA19–9 antigen, 430  
 CACNA1C mutation, 141  
 CADASIL. *See* Cerebral arteriopathy autosomal dominant with subcortical infarcts and leukoencephalopathy (CADASIL)

- Cadherins, 21  
 CAH. *See* Congenital adrenal hyperplasia (CAH)  
 CAIS. *See* Complete androgen insensitivity syndrome (CAIS)  
 Cajal bodies, speckles and pre-mRNA processing, 29–30  
 CAKUT. *See* Congenital anomalies of the kidneys and urinary tract (CAKUT)  
 Calcium  
   calcium-induced calcium release, 136–137  
   homeostasis, 475  
   abnormalities, 475  
   sparks, 145  
 Calmodulin (CaM), 145  
 Calsequestrin, 145  
 CaM. *See* Calmodulin (CaM)  
 cAMP. *See* 3',5'-Cyclic monophosphate (cAMP)  
 Cancer, 47, 49–51, 525  
   breast, 73, 424–425  
   drug development, 499  
   genetic imprinting and, 426–427  
   immunogenetics and, 426  
   precision medicine, 523  
   RNA interference and, 426  
 Cancer antigen (CA), 430  
 Cancer family syndromes, 425  
   based on clinical phenotypes, 425  
   genetic imprinting and cancer, 426–427  
   immunogenetics and cancer, 426  
   MEN, 425  
   RAS–MAPK syndromes, 425  
   RNA interference and cancer, 426  
   VHL disease, 426  
 Cancer paradigm  
   histology agnostic treatment, 499–500  
   making common cancer rare, 501  
   molecularly targeted therapy limitations in cancer, 500–501  
   oncogene addiction, 497–499  
   synthetic lethality, 499  
 Cancer susceptibility  
   cancer family syndromes, 425  
   complex cancer genomics, 427  
   diagnostic and predictive genetic testing for cancer, 431–434  
   DNA repair genes, 424–425  
   genetic counseling for inherited, 430–431  
   inherited and familial cancer, 423  
   inherited susceptibility to leukemia, 427–430  
   oncogenes and tumor suppressor genes, 424  
   tumor markers in circulating blood, 430  
*Candida* species, 415–416  
 Candidate genes, 359  
   association studies, 438  
   studies, 357–359, 403–404  
 CAP. *See* Community-acquired pneumonia (CAP)  
 Capillary electrophoresis (CE), 51  
 Carbamazepine (CBZ), 525  
   HLA testing for prediction  
   Stevens–Johnson syndrome with uses, 527  
 Carbon dioxide (CO<sub>2</sub>), 43  
 Carcinoembryonic antigen, 430  
 CARD9 gene, 377  
 Cardiac action potential, 135  
 Cardiac biopsy, 160  
 Cardiac dyssynchrony, 160  
 Cardiac genetics multidisciplinary team, 163  
 Cardiac magnetic resonance imaging (CMR imaging), 159–160  
   tissue characterization, 159f  
 Cardiac myosin, 320–321  
 Cardiac resynchronization therapy (CRT), 156  
 Cardiac structural abnormality, 156  
 Cardioembolic stroke, 191–193  
 Cardiomyocytes, 136–137  
 Cardiomyopathy. *See also* Heart failure phenotype, 162–163  
   screen, 155  
 Cardiotoxic drugs, 155  
 Cardiovascular developmental disorder, 121.  
   *See also* Heart failure  
   advances in technology, 122–123  
   chromosomal aneuploidy and structural heart disease, 123–124  
   epigenome and congenital heart disease, 128–129  
   future considerations, 129  
   historical overview, 121  
   noncardiac congenital anomalies in CHD, 128  
   noncoding regulatory genome in CHD, 126–128  
   normal development of heart, 121–122  
   single-gene disorders, 124–126, 125t  
 Cardiovascular disease risk assessment (CVD risk assessment), 249–251  
 Carpenter–Waziri syndrome, 90–91  
 CARS. *See* Compensatory antiinflammatory syndrome (CARS)  
 Cascade testing, 133, 255, 430–431  
 Case probands, 353  
 CASQ2, 145  
 Catecholaminergic polymorphic ventricular tachycardia (CPVT), 134, 145–146, 148–149  
   clinical risk assessment and therapy in, 148–149  
   generation and propagation of arrhythmia in, 145  
   molecular risk stratification in, 145–146  
 Causative genome variants, 378  
 CAV3 protein, 141  
 Caveolae, 17–18, 38  
 CBZ. *See* Carbamazepine (CBZ)  
 CC. *See* Cervical carcinomas (CC)  
 CC chemokine receptor-2 (CCR2), 404  
 CCAAT enhancer-binding protein (C/EBP), 348  
 CCAAT-enhancer binding protein (CEBPA), 427–428  
 CCL2. *See* Chemokine (C-C motif) ligand 2 (CCL2)  
 CCR2. *See* CC chemokine receptor-2 (CCR2)  
 CD. *See* Cluster of differentiation (CD); Crohn's disease (CD)  
 CD14, 357–358  
 CD234. *See* Fy glycoprotein  
 CD45 gene polymorphism C77G, 407–408  
 CDC. *See* Complement-dependent cytotoxicity (CDC)  
 CDC 73. *See* Cell division cycle 73 (CDC 73)  
 CDGP. *See* Constitutional delay in growth and puberty (CDGP)  
 CDK gene, 224–225  
 CDKI. *See* Cyclin-dependent kinase inhibitors (CDKI)  
 CDKN1A gene, 41  
 CDKN1B gene, 225  
 CDKN2B antisense RNA 1 gene (CDKN2B-AS1), 195  
 CDKs. *See* Cyclin-dependent kinases (CDKs)  
 CDON genes, 92  
 CE. *See* Capillary electrophoresis (CE)  
 CEBPA. *See* CCAAT-enhancer binding protein (CEBPA)  
 Cell  
   adhesion molecules, 21, 92  
   cortex, 20–21  
   cycle, 39–41, 40f  
   death, 44, 193–194, 452–453  
   junctions, 19–21  
   nucleus and gene expression, 23–32, 24f  
   Cajal bodies, speckles and pre-mRNA processing, 29–30  
   chromosome territories, gene transcription and nuclear lamina, 23–29  
   genes and genetic diseases associated with organelle dysfunction, 25f  
   nuclear envelope and mRNA quality control, 30–32  
   nucleolus, 30  
   signaling, 18–19  
 Cell division cycle 73 (CDC 73), 224–225, 226f  
 Cell-free circulating fetal DNA, 332  
 Cell-free circulating tumor DNA (ctDNA), 501  
 Cellular glucose transport, 268  
 Cellular rejection in HLA, 319  
 Central nervous system (CNS), 447  
 Centrioles, 39–40  
 Centrosomes, 22, 39, 41  
 Cerebellar degeneration, 476, 477t  
 Cerebral amyloid angiopathy, 201  
 Cerebral arteriopathy autosomal dominant with subcortical infarcts and leukoencephalopathy (CADASIL), 190  
 Cerebrospinal fluid (CSF), 457  
 Ceritinib, 498  
 Cervical cancer, 408  
   host genetic susceptibility to HPV infection and development of, 408  
 Cervical carcinomas (CC), 404  
 CETP. *See* Cholesteryl ester transfer protein (CETP)  
 CF. *See* Cystic fibrosis (CF)  
 CFH gene, 404  
 CFTR. *See* Cystic fibrosis transmembrane conductance regulator (CFTR)



- CGD. *See* Chronic granulomatous disease (CGD)
- CGH. *See* Comparative genomic hybridization (CGH)
- CH. *See* Congenital hypopituitarism (CH)
- Channelopathies, 133  
Brugada syndrome, 146–147  
catecholaminergic polymorphic ventricular tachycardia (CPVT), 145–146  
clinical cases, 133–135  
mechanisms, 137  
molecular biology and pathophysiology, 137–144  
molecular systems underpinning clinical scenario, 135–137  
short QT syndrome (SQTS), 144  
targeted molecular diagnosis and therapy, 147–149
- Chaperones, 36
- Charcot–Marie–Tooth disease type 2B1, 94–95
- CHARGE syndrome, 89, 98*t*
- CHD. *See* Congenital heart disease (CHD); Coronary heart disease (CHD)
- CHD7 genes, 89, 125–126
- Chemokine (C-C motif) ligand 2 (CCL2), 203
- Chemokine receptor (CCR)  
CCR2, 408  
CCR5, 403–404  
genetic variants, 404–405
- CHH. *See* Congenital hypogonadotropic hypogonadism (CHH)
- Child health-care center, 75
- Childhood overgrowth syndromes, 426
- Chlamydia pneumoniae*, 413–414
- Chlamydia psittaci*, 413
- Chlamydia trachomatis*, 403, 413–414
- Choice of test, 59–62
- Cholesteryl ester transfer protein (CETP), 246–247
- CHOPS syndrome, 98*t*
- Chromatin, 12
- Chromatin immunoprecipitation, 373
- Chromatin remodeling, 98*t*
- Chromosomal aneuploidy and structural heart disease, 123–124  
CHD and CNVs, 123–124  
chromosomal syndromes, 124*t*
- Chromosomal microarray (CMA), 57
- Chromosomes, 4, 5*f*  
breakage syndromes, 429  
chromosome 1q42–q43, 145  
chromosome 2 (q12.3), 412  
chromosome 5q, 357  
chromosome 6p, 357  
chromosome 8q12, 125–126  
chromosome 11, 357  
chromosome 12q, 357  
microtubules, 41  
territories, 23–29
- Chronic alcohol-induced inflammation, 391
- Chronic granulomatous disease (CGD), 416
- Chronic HBV infection, 315
- Chronic hypertransfusion therapy, 328–329
- Chronic hypocalcemia, 227–228  
clinical case, 227–228  
management of patients with PHP, 228  
molecular systems, 227–228
- Chronic hypophosphatemia, 228–229  
clinical case, 228–229  
management of patients with XLH, 229  
molecular systems, 229
- Chronic inflammation, 389–390  
age-associated, 392–393  
microvascular environment theory, 389*f*  
molecular diagnosis and treatment, 393–400  
molecular pathology, 391  
of acute inflammation, 390–391  
molecular therapy  
for acute inflammatory diseases, 400  
for chronic inflammatory diseases, 400
- Chronic inflammatory connective tissue diseases, 394–400  
antinuclear antibodies in, 399*t*  
RA, 395–400
- Chronic inflammatory diseases, 393–394
- Chronic kidney disease (CKD), 438–440  
causes, 439  
establishing molecular genetic diagnosis, 439  
clinical heterogeneity, 439  
genetic heterogeneity, 439  
genetic variants and, 440  
personalized pharmacotherapy in, 441  
targeted gene panel analysis in, 440–441  
understanding pathogenesis, 439–440
- Chronic low-grade inflammation, 393
- Chronic lymphocytic leukemia (CLL), 427
- Chronic mucocutaneous candidiasis (CMC), 415
- Chronic myeloid leukemia (CML), 428–429, 497
- Chronic neurodegeneration  
molecular diagnostics application in neurodegeneration, 456–458  
molecular pathology of neurodegenerative diseases, 453–455  
cellular vulnerability, 455*f*  
vulnerability patterns in neurodegenerative diseases, 454*f*  
neurodegenerative disease clinical case studies and molecular systems, 448–453  
AD, 449–450  
FTD, 452–453  
PD, 451–452
- Chronic Q fever, 414–415
- Chudley–Lowry syndrome, 90–91
- Chylomicron, 245
- Ciliopathies, 95–96  
revealing overlapping phenotype and genetic heterogeneity, 96*t*
- Cinacalcet, 226
- Circular RNAs (circRNAs), 126–128
- Circulating blood, tumor markers in, 430
- Citrullination, 395
- CKD. *See* Chronic kidney disease (CKD)
- Clathrin-coated vesicles, 37
- Claudin, 20
- Cleavage–polyadenylation process, 329
- Clinical applications of metabolomics, 47–51
- Clinical dysmorphology, 89
- Clinical metabolomics, 47, 51*t*  
applications, 47–51  
in cancer and human diseases, 49–50  
inborn errors of metabolism, 47–49  
databases, 51–52  
techniques in, 51–52
- Clinical molecular nephrology  
AKI, 437–438  
CKD, 438–440  
clinical renal genomic medicine, 440–441  
molecular basis of kidney transplantation, 442
- Clinical Pharmacogenetics Implementation Consortium (CPIC), 529–530
- Clinical Practice Guideline, 83
- Clinical renal genomic medicine, 440–441  
genetic counseling, 441  
personalized pharmacotherapy in CKD, 441  
targeted clinical management, 441  
targeted clinical surveillance, 441  
targeted gene panel analysis in CKD, 440–441
- Clinical testing approach, 300–301
- Clinical whole exome sequencing, 294
- ClinVar, 67
- CLL. *See* Chronic lymphocytic leukemia (CLL)
- Clonazepam, 471
- Clopidogrel, 528
- Clozapine, 471
- Cluster of differentiation (CD), 404
- Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 system (CRISPR/Cas protein 9 system), 516–517
- CMA. *See* Chromosomal microarray (CMA)
- CMC. *See* Chronic mucocutaneous candidiasis (CMC)
- CML. *See* Chronic myeloid leukemia (CML)
- CMR imaging. *See* Cardiac magnetic resonance imaging (CMR imaging)
- CMV. *See* Cytomegalovirus (CMV)
- CNP acts. *See* C-type natriuretic peptide acts (CNP acts)
- CNS. *See* Central nervous system (CNS)
- CNVs. *See* Copy-number variants/ variations (CNVs)
- Coagulation, 341–343, 342*f*, 393  
genetic basis  
of bleeding disorders, 345–348  
of thrombosis, 343–345  
inherited defects of platelets, 349  
structural defects of vascular system, 348
- Coat protein complexes I (COPI), 36
- Coat protein complexes II (COPII), 36
- Coding RNA (cRNA), 12
- Codominant inheritance, 354
- Codons, 4–6
- Coenzyme A synthetase protein-associated neurodegeneration (COASY protein-associated neurodegeneration), 482
- Coffin–Siris syndrome, 98*t*



- Cognitive function among CNVs carriers, 489
- Cognitive impairment, 452
- Cohorts for heart and aging research in genomic epidemiology (CHARGE) association, 125–126
- syndrome, 89
- COL4A1. *See* Collagen, type IV, alpha 1 gene (COL4A1)
- Collagen, 320–321
- Collagen, type IV, alpha 1 gene (COL4A1), 190
- Colon cancer, 73
- Colonic epithelial IRGM gene, 379
- Colonoscopy, 50
- Colorectal cancer (CRC), 50, 423, 425
- Common hypercholesterolemia. *See* Hypercholesterolemia—polygenic
- Common variant—common disease, 359
- Community-acquired pneumonia (CAP), 413
- Comparative genomic hybridization (CGH), 301
- Comparative genomics, 6–7
- Compartment lumen pH, 37
- Compensatory antiinflammatory syndrome (CARs), 390
- Complement activation in HLA, 318–319
- Complement component 4 genes (C4 genes), 490–491
- Complement-dependent cytotoxicity (CDC), 318
- Complete androgen insensitivity syndrome (CAIS), 112–113
- Complex antigen–antibody interaction, 395
- Complex disease, 169, 353
- Computerized dysmorphology analysis, 90
- Computerized tomography (CT), 157
- Congenital adrenal hyperplasia (CAH), 111, 171, 178–180, 233
- 17 $\alpha$ -hydroxylase deficiency, 179–180
- 11 $\beta$ -hydroxylase deficiency, 178–179
- case report, 178–179
- classic, 233
- clinical case, 233–234
- management, 234
- molecular systems, 234
- diagnosis, 181
- management, 181
- Congenital anomalies of the kidneys and urinary tract (CAKUT), 439
- Congenital cyanosis, 336
- Congenital gamma-chain molecular defect, 345
- Congenital heart disease (CHD), 121–124
- epigenome and, 128–129
- noncardiac congenital anomalies in, 128
- noncoding regulatory genome in, 126–128
- Congenital hypogonadotropic hypogonadism (CHH), 217–221
- clinical case, 219–221
- correlation of underlying gene mutation, 220 $t$
- molecular systems, 220–221
- Congenital hypopituitarism (CH), 217–221
- clinical case, 217–218
- genetic and clinical characteristics, 219 $t$
- management, 218
- molecular systems, 218
- Congenital LQTS, 138, 138 $t$
- Connexin family, 20
- Constitutional delay in growth and puberty (CDGP), 219–220
- Constitutional MMR deficiency syndrome, 425
- Constitutive transcription, 28
- Control probands, 353
- COPI. *See* Coat protein complexes I (COPI)
- Copy-number variants/variations (CNVs), 9, 83, 123–124, 192, 296–298, 372, 440, 487–489, 493
- cognitive function among carriers, 489
- creation, 488 $f$
- penetrance, 489–490, 490 $t$
- in schizophrenia patients and controls, 488 $t$
- CoQ10 deficiency, 441
- Coreceptor and ligand variants in HIV disease progression, 404
- Cornelia/Brachman de lange syndrome, 98 $t$
- Coronary heart disease (CHD), 247
- Corrected QT interval (QT $_c$ ), 147
- Corticosteroids receptor antagonists, 363
- Cowden syndrome (CS), 424
- Coxiella burnetii*, 414–415
- CpG. *See* Cytosine-phosphate-guanosine (CpG)
- CPIC. *See* Clinical Pharmacogenetics Implementation Consortium (CPIC)
- CPVT. *See* Catecholaminergic polymorphic ventricular tachycardia (CPVT)
- CRC. *See* Colorectal cancer (CRC)
- CRISPR RNA (crRNA), 516–517
- CRISPR/Cas protein 9 system. *See* Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 system (CRISPR/Cas protein 9 system)
- Crizotinib, 498, 523
- cRNA. *See* Coding RNA (cRNA)
- Crohn's disease (CD), 359, 367
- with UC, 368 $t$
- Cross presentation, 311–312
- CRP. *See* C-reactive protein (CRP)
- crRNA. *See* CRISPR RNA (crRNA)
- CRT. *See* Cardiac resynchronization therapy (CRT)
- Cryptic splice sites, 329
- Cryptococcosis, 417
- Cryptococcus*, 362
- C. gattii*, 416–417
- C. neoformans*, 416–417
- CS. *See* Cowden syndrome (CS)
- CSF. *See* Cerebrospinal fluid (CSF)
- CST3. *See* Cystatin C (CST3)
- CT. *See* Computerized tomography (CT)
- ctDNA. *See* Cell-free circulating tumor DNA (ctDNA)
- CTLA4. *See* Cytotoxic T-lymphocyte-associated protein-4 (CTLA4)
- CUL3 gene, 173–175
- CVD risk assessment. *See* Cardiovascular disease risk assessment (CVD risk assessment)
- Cyanotic CHD, 126
- Cyclic GMP, 19
- 3',5'-Cyclic monophosphate (cAMP), 19
- Cyclin proteins, 225
- Cyclin-dependent kinase inhibitors (CDKI), 225
- Cyclin-dependent kinases (CDKs), 40, 225
- CYD-TDV. *See* Dengvaxia (CYD-TDV)
- CYP enzymes. *See* Cytochrome P450 enzymes (CYP enzymes)
- CYP11B1 gene, 178–179
- CYP17A1 gene, 180
- CYP21A2 gene, 114
- CYP2C9 gene, 526–527
- testing for CYP2C19 activity prior, 528
- CYP2D6 gene, 524
- Cystatin C (CST3), 190
- Cysteine-rich, transmembrane module (CYSTM), 457
- Cystic fibrosis (CF), 10, 101–102
- Cystic fibrosis transmembrane conductance regulator (CFTR), 101–102
- CYSTM. *See* Cysteine-rich, transmembrane module (CYSTM)
- Cytochrome C, 44
- Cytochrome enzymes, 300
- Cytochrome P450 enzymes (CYP enzymes), 300, 524
- Cytokines, 319, 390
- cytokine-mediated process, 373
- polymorphism and dengue, 410–411
- Cytomegalovirus (CMV), 392–393, 513
- Cytosine (C), 4
- Cytosine-phosphate-guanosine (CpG), 516
- dinucleotide, 128
- islands, 128
- shores, 128
- Cytoskeleton, 21–22
- Cytosol, 36
- Cytotoxic T-lymphocyte-associated protein-4 (CTLA4), 272
- D**
- D-loop. *See* Displacement loop (D-loop)
- Dabrafenib, 497–499
- Danish twin registry, 355
- Dapsone, 524
- Daughter centriole, 40
- DBP. *See* Diastolic BP (DBP)
- DBS. *See* Deep brain stimulation (DBS)
- DCs. *See* Dendritic cells (DCs)
- DCX X-linked gene, 22
- DDR. *See* DNA-damage response (DDR)
- De novo mutations, 129
- De novo SNVs, 492
- De novo status and segregation data, 67–68
- Decoy oligonucleotides, 516
- Dectin-1 gene, 377
- Deep brain stimulation (DBS), 464, 471
- Deep sequencing, 374
- Defective MMR (dMMR), 500
- Defective regulation of complement pathway, 393
- Defibrillators, 158
- Deformation, 89

- 19 Del. *See* Deletion in exon 19 (19 Del)  
 Delayed cerebral ischemia, 204  
 Deletion in exon 19 (19 Del), 498–499  
 Deletion polymorphism, 372  
 $\delta$ -aminolevulinic acid synthase-2 (ALAS-2), 336  
 Dendritic cells (DCs), 379, 406  
 Dengue and major histocompatibility complex antigens, 410  
 Dengue fever (DF), 410  
 Dengue hemorrhagic fever (DHF), 410  
 Dengue infections (DI), 410  
 Dengue viral infection, 410–411  
   cytokine polymorphism and dengue, 410–411  
   dengue and major histocompatibility complex antigens, 410  
 Dengue virus (DV), 410  
 Dengvaxia (CYD-TDV), 410–411  
 Dense and alpha granule deficiency, 349  
 Dentarubral–pallidolusian atrophy (DRPLA), 483  
 Denys–Drash syndrome, 111  
 11-Deoxycorticosterone, 178, 181  
 Deoxyribonuclease I hypersensitive sites (DHSs), 378  
 Deoxyribonucleic acid (DNA), 4, 6, 89, 341–344  
   analysis, 115  
   diagnostics, 83  
   DNA-based tests, 341  
   elements, 28  
   and gene content of reference human genome, 8*t*  
   methylation, 128–129, 316  
   recombinant HPV vaccine, 507  
   repair genes, 424–425  
     breast and ovarian cancer, 424–425  
     colorectal cancer, 425  
   technologies, 505–506  
 Depolarization theories, 146–147  
 Depression, 135, 471, 480–481  
 Desmoplakin, 22  
 Desmosomes, 21  
 Developmental delay, 61  
 DF. *See* Dengue fever (DF)  
 DGKE. *See* Diacylglycerol kinase epsilon (DGKE)  
 DHF. *See* Dengue hemorrhagic fever (DHF)  
 DHR. *See* Drug hypersensitive reaction (DHR)  
 DHSs. *See* Deoxyribonuclease I hypersensitive sites (DHSs)  
 DHT. *See* Dihydrotestosterone (DHT)  
 DI. *See* Dengue infections (DI)  
 Diabetes mellitus (DM), 267, 271*t*, 439, 505  
   diagnosis, 274–276  
     blood glucose parameters, 275–276  
     clinical manifestations, 274–275  
     genetic testing, 276  
     HbA1C correlation, 276  
   disorders of glycemic regulation, 270–274  
   inherited monogenic diabetes mellitus, 283–284  
   management, 285–286  
   molecular basis of glycemic homeostasis, 267–270  
   overweight, obesity, and, 276–281  
     constitutional and medical obesity, 278  
     genetic/genomic factors, 279–281  
     neurobiological factors, 276–277  
     nutritional factors, 277  
     vitamin D and, 281–282  
 Diacylglycerol, 19  
 Diacylglycerol kinase epsilon (DGKE), 438  
 Diagnostic genetic testing, 437  
 Diagnostic odyssey, 86, 91  
 Diagnostic testing for cancer, 431–434  
 Diastolic BP (DBP), 169  
 DICER1 gene, 426  
 Diet/nutritional supplements, 285  
 Dietary intake, 370  
 DiGeorge syndrome. *See* 22q11.2 deletion  
 Dihydrotestosterone (DHT), 242  
 Diploid, 5*f*, 6  
 Direct genotyping variants, 371–372  
 Direct pathway, 317–318  
 Direct PT G20210A mutation testing, 344  
 Diseases, 156–157  
   allele, 371  
   disease-causing gene discovery, 461  
   disease-causing mutations, 371  
   heterogeneity, 361  
 Disorders of sex development (DSD), 107  
   diagnosis, 108*f*  
   gender assignment, 116–117  
   genes, 116  
   gonadal cancer risk, 117  
   initial assessment, 108*f*  
   investigations, 115–116  
   sex chromosome, 108–110  
     46, XX, 113–115  
     XY disorders of sexual differentiation, 110–113  
 DISP1 genes, 92  
 Displacement loop (D-loop), 11  
 Disruption, 89  
 Dizygotic twins (DZ twins), 354, 369–370  
 DJ1/PARK7, 463  
 DLL1 genes, 92  
 DM. *See* Diabetes mellitus (DM)  
 DMD. *See* Duchenne muscular dystrophy (DMD)  
 dMMR. *See* Defective MMR (dMMR)  
 DMRT1 genes, 111  
 DMRT2 genes, 111  
 DNA. *See* Deoxyribonucleic acid (DNA)  
 DNA methyltransferase (DNMT), 373  
 DNA-based testing, 521–522  
 DNA-damage response (DDR), 499  
 DNMT. *See* DNA methyltransferase (DNMT)  
 Dominant  $\beta$  thalassemia, 330  
 Donohue syndrome, 284  
 Donor selection, 442  
 Donor-specific antibodies (DSAs), 318  
 Doose syndrome, 293  
 Dopa-responsive dystonia (DYT5), 469  
 Double diabetes, 270  
 Double-strand breaks (DSBs), 517  
 Down syndrome (DS), 427  
 Dravet syndrome, 293, 303  
 DRPLA. *See* Dentarubral–pallidolusian atrophy (DRPLA)  
 Drug hypersensitive reaction (DHR), 313  
 Drug response, 300  
   validation of genetic/genomic information for, 528  
 Drug safety, 525  
 Drug-induced dyskinesias, 464  
 Drug-induced hypersensitivities, HLA and, 313  
 Drug-induced long QT, 143  
 Drug-metabolizing enzymes, 524  
 Drug-refractory epilepsy, 300–301  
 DS. *See* Down syndrome (DS)  
 DSAs. *See* Donor-specific antibodies (DSAs)  
 DSBs. *See* Double-strand breaks (DSBs)  
 DSD. *See* Disorders of sex development (DSD)  
 DSP gene coding, 22  
 Duchenne muscular dystrophy (DMD), 99–100, 513  
 Duffy, 417  
 Duffy antigen/chemokine receptor (DARC). *See* Duffy  
 Dutch Lipid Clinic Networks criteria, 254  
 DV. *See* Dengue virus (DV)  
 Dysfibrinogenemia, 348  
 Dysmorphology, 89  
   clinical cases and molecular basis, 91–97  
     epigenetic mechanisms, 96–97  
     genetic heterogeneity, 92–96  
     spliceopathies, 97  
     transcriptomopathies, 96–97  
   databases, 89  
   limitations of “clinical” dysmorphology and tools, 89–90  
   molecular, 90–91  
     diagnosis and therapy, 98–103  
     and molecular medicine, 91  
   molecular diagnosis and therapy, 98–103  
 Dystonia, 471  
   clinical and genetic description of dystonic disorders, 466*t*  
   clinical characteristics, 464–465  
   diagnosis, 465–470  
     DYT1, 469  
     DYT10, 470  
     DYT11, 470  
     DYT5, 469  
     DYT6, 469–470  
   genetics, 465  
   pathophysiology, 465  
   targeted molecular diagnosis and therapy of, 465  
   therapy, 470–471  
     botulinum toxin, 471  
     DBS, 471  
     oral medication, 471  
     physiotherapy, 471  
 Dystonic cerebral palsy, 464–465  
 Dystrophic epidermolysis bullosa, 103  
 Dystrophin, 21  
 DYT1, 465, 469  
 DYT5. *See* Dopa-responsive dystonia (DYT5)

- DYT6, 469–470  
 DYT10, 470  
 DYT11, 465, 470  
 DZ twins. *See* Dizygotic twins (DZ twins)
- E**  
 4E-BP1. *See* Eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1)  
 e-QTL. *See* Expression quantitative trait loci (e-QTL)  
 E1784K, 143  
 Early endosome, 37  
 EBV. *See* Epstein–Barr virus (EBV)  
 ECG. *See* Electrocardiography/  
 electrocardiogram (ECG)  
 Echocardiography, 133, 158–159  
 laboratory investigations in, 156*t*  
 Eculizumab, 438  
 EDS. *See* Ehlers–Danlos syndrome (EDS)  
 EEG. *See* Electroencephalogram/  
 electroencephalography (EEG)  
 EGFR. *See* Epidermal growth factor receptor (EGFR)  
 Ehlers–Danlos syndrome (EDS), 348  
 eIF2A. *See* Eukaryotic translation initiation factor 2A (eIF2A)  
 eIF4E. *See* Eukaryotic translation initiation factor 4E (eIF4E)  
 eIF4G. *See* Eukaryotic translation initiation factor 4G (eIF4G)  
 EJC. *See* Exon junction complex (EJC)  
 Elastin gene (*ELN*), 205  
 Electrocardiography/electrocardiogram (ECG), 133, 156–157  
 ambulatory, 156–157  
 resting, 156  
 Electroclinical syndromes, 294  
 Electroencephalogram/  
 electroencephalography (EEG), 289, 291*f*, 296  
 normal waking, 297*f*  
 Electronmicroscopical studies, 199  
 Elliptocytosis, 417  
*ELN*. *See* Elastin gene (*ELN*)  
 Elongation of protein synthesis, 34  
 EMA. *See* European Medicines Agency (EMA)  
 ENaC. *See* Epithelial sodium channel (ENaC)  
 End-stage renal disease (ESRD), 437  
 Endocrine signaling, 18–19  
 Endocrinology, clinical molecular, 217  
 apparent mineralocorticoid excess, 233–236  
 CAH, 233–236  
 CH, 217–221  
 CHH, 217–221  
 chronic hypocalcemia and  
 hypophosphatemia, 227–229  
 HPT, 222–226  
 multiple endocrine neoplasia syndromes, 222–226  
 PA, 230–232  
 pituitary adenoma, 217–221  
 primary gonadal failure and androgen resistance/insensitivity syndrome, 240–242  
 renal tubular defects, 233–236  
 Endocytic pathways, 34–38, 35*f*  
 Endogenous damage signals, 392  
 Endoplasmic reticulum (ER), 19, 475  
 Endoplasmic reticulum aminopeptidase (ERAP1), 314  
 Endothelial cells, 204  
 Endothelial nitric oxide synthase (eNOS), 127–128  
 Endothelium, direct activation of, 319  
 Endotoxemic shock, 393  
 Energy  
 energy-hungry systems, 298  
 expenditure, 79  
 intake, 79  
 metabolism, 298  
 production, 42–44  
 Engerix-B, 507  
 Enhancer polymorphisms-mediated expression variability, 316  
 eNOS. *See* Endothelial nitric oxide synthase (eNOS)  
 Enteropathogens, genetics of susceptibility to, 419–420  
 host receptors, 419  
 innate immune  
 genes associated with increased susceptibility, 419  
 response and cellular injury, 419–420  
 Environmental factors, 92, 189, 201, 253–254  
 obesogenic factors, 73  
 risk factors, 171  
 Environome, 369–370  
 Enzyme-replacement therapy (ERT), 99, 508  
 Epidermal growth factor receptor (EGFR), 498, 523  
 Epigenetics, 128, 355, 373  
 mechanisms, 96–97  
 phenotypic spectrum of filaminopathies, 97*t*  
 regulatory enzymes, 373  
 Epigenome, 12, 373  
 and congenital heart disease, 128–129  
 DNA methylation, 128–129  
 histone modification, 129  
 Epilepsy, 289–291, 299  
 diagnosis of progressive myoclonic, 303*t*  
 evidence for genetic basis, 291–292  
 familial focal, 293  
 genes, 302  
 genetic architecture, 292–298  
 CNVs, 296–298  
 GWAS, 293–294  
 karyotypic abnormalities, 296  
 linkage studies in families, 293  
 noncoding variants, 295–296  
 rare variants in epileptic encephalopathies, 294  
 idiopathic generalized, 291–292  
 mitochondrial, 298–299  
 molecular genetic testing strategies for, 300–304  
 pharmacogenetics of, 299–300  
 phenotype, 292–293  
 progressive myoclonic, 299  
 rare coding sequence variants in, 294–295  
 syndrome, 290  
 Epileptic encephalopathies, 290, 294  
 genetic, 294  
 rare variants in, 294  
 Epistasis, 358–359  
 Epithelial cell organelles and structures, 18*f*  
 Epithelial sodium channel (ENaC), 170  
 Epsilon-sarcoglycan gene (*SGCE* gene), 470  
 ( $\epsilon\gamma\delta\beta$ )<sup>0</sup> thalassemia, 330  
 Epstein–Barr infection, host genetics of, 409–410  
 Epstein–Barr virus (EBV), 409  
 ER. *See* Endoplasmic reticulum (ER)  
 ERAP1. *See* Endoplasmic reticulum aminopeptidase (ERAP1)  
 ERBB2 genotype, 526  
 Erlotinib, 523  
 ERT. *See* Enzyme-replacement therapy (ERT)  
 Erythrocytes, 417  
 Erythroid heme biosynthesis, defects of, 336  
 ESC. *See* European Society of Cardiology (ESC)  
*Escherichia coli*, 505–506  
 ESRD. *See* End-stage renal disease (ESRD)  
 Ethical issues in genomic medicine, 532  
 Etiology  
 genetic, 89  
 of heart failure, 153–154  
 of intracerebral hemorrhage, 199*f*  
 of ischemic stroke, 193*f*  
 of subarachnoid hemorrhage, 202*f*  
 Eukaryotes, 4  
 gene expression, 23*b*  
 Eukaryotic  
 cells, 17  
 genomes, 23–24  
 release factor, 34  
 Eukaryotic translation initiation factor 2A (eIF2A), 32–33  
 Eukaryotic translation initiation factor 4E (eIF4E), 32  
 Eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1), 32  
 Eukaryotic translation initiation factor 4G (eIF4G), 32  
 European Medicines Agency (EMA), 532  
 European Society of Cardiology (ESC), 153  
 Evidence-based clinical diagnosis, 394  
 Evidence-based medicine, 537–539, 538*f*  
 Excitation–contraction coupling, 135–137  
 Exercise testing, 160  
 Exercise tolerance test, 133  
 Exome, 61  
 panel, 301, 303  
 sequencing, 409, 492  
 Exon junction complex (EJC), 30  
 Expanded panel, 61  
 Exposure or overproduction of androgens, 114–115  
 Expression quantitative trait loci (e-QTL), 311, 371–372  
 External stimuli, 353  
 Extracorporeal LDL removal, 259–260  
 Extramedullary hematopoiesis, 328–329  
 Ezetimibe, 259

- F**
- Fabry's syndrome, 508
- Factor II G20210A. *See* Prothrombin allele G20210A
- Factor IX deficiency (FIX deficiency). *See* Hemophilia B
- Factor V Leiden (FVL), 344–345
- Factor VIII deficiency. *See* Hemophilia A
- FADH2. *See* Flavin adenine dinucleotide (FADH2)
- Faecalibacterium prausnitzii*, 370
- FAHN. *See* Fatty acid hydroxylase-associated neurodegeneration (FAHN)
- Familial adenomatous polyposis (FAP), 21, 423–424
- Familial cancer, 423
- Familial clustering, 425
- Familial combined hyperlipidemia (FCH), 249
- Familial erythroleukemia, 428
- Familial focal epilepsy, 293
- Familial hyperaldosteronism (FH), 181–183, 230
- case report, 181
- diagnosis, 182
- genetics, 182
- management, 182–183
- pathophysiology, 182
- Familial hypercholesterolemia (FH), 249, 252
- Familial hypocalciuric hypercalcemia (FHH), 222
- genetics and prevalence, 225*t*
- subtype 1, 224–225
- Familial renal glycosuria, 275–276
- FAP. *See* Familial adenomatous polyposis (FAP)
- Farber/spinal muscular atrophy with progressive myoclonic epilepsy (SMA-PME), 102–103
- Fas, 408
- Fas ligand (FasL), 408
- FASTQ file, 65
- Fatty acid hydroxylase-associated neurodegeneration (FAHN), 482
- Fatty acid oxidation, 43
- FCH. *See* Familial combined hyperlipidemia (FCH)
- FCXM. *See* Flow crossmatch (FCXM)
- FDA. *See* Federal Drug Agency (FDA)
- Febrile nonhemolytic transfusion reactions (FNHTR), 322
- Federal Drug Agency (FDA), 513
- Fetal erythrocytes, 328
- FG syndrome, 98*t*
- FGF. *See* Fibroblast growth factor (FGF)
- FGFR3 decoy receptor, 100
- FH. *See* Familial hyperaldosteronism (FH); Familial hypercholesterolemia (FH); Fumarate hydratase (FH)
- FHH. *See* Familial hypocalciuric hypercalcemia (FHH)
- Fibrates, 260
- Fibrinolysis, inherited deficiency of, 345
- Fibroblast growth factor (FGF), 220–221
- FGF8 genes, 92
- Fibroblast growth factor 3 (FGFR3) gene, 100
- Fibrous protein genes, 348
- Filaminopathies, 96
- First generation sequencing, 58
- FISH. *See* Fluorescent in situ hybridization (FISH)
- Flavin adenine dinucleotide (FADH2), 43
- Flavivirus*, 410
- Flexible filament actin (F-actin), 21
- 18-Florbetaben PET images, 450
- Flow crossmatch (FCXM), 319
- Fluorescent in situ hybridization (FISH), 91
- Fluorescent reversible-terminator nucleotides, 59
- FMRP. *See* Fragile X mental retardation protein (FMRP)
- FNHTR. *See* Febrile nonhemolytic transfusion reactions (FNHTR)
- Focal segmental glomerulosclerosis (FSGS), 440
- Follicle-stimulating hormone (FSH), 217
- Follicular stimulatory hormone (FSH), 220
- Forkhead box L2 (FOXL2), 110–111
- FOXH1 genes, 92
- FOXL2. *See* Forkhead box L2 (FOXL2)
- Fragile X mental retardation protein (FMRP), 490–491
- Frameshift mutation, 66, 81, 102, 176, 330
- C-terminal, 190
- in MC4R, 81
- Framingham risk equation, 528
- Frataxin* gene (FXN gene), 476
- Fred Sanger's chain termination method, 58
- Freidreich's ataxia, 476
- Frontotemporal dementia (FTD), 452–453
- FSGS. *See* Focal segmental glomerulosclerosis (FSGS)
- FSH. *See* Follicle-stimulating hormone (FSH); Follicular stimulatory hormone (FSH)
- FSHR* genes, 113
- FTD. *See* Frontotemporal dementia (FTD)
- FtL*. *See* L chain of ferritin (*FtL*)
- Fucosyltransferase 2 deficiency (FUT2 deficiency), 379–380, 403–404
- Fumarate hydratase (FH), 426
- Functional genomics, 7, 12–13
- Fungal infections, host genetic susceptibility and protection from, 415–417
- Aspergillosis*, 416
- genetic variability of host and susceptibility to invasive aspergillosis, 416
- Candida*, 415–416
- Cryptococcus gattii*, 416–417
- Cryptococcus neoformans*, 416–417
- Fusobacterium*, 362
- FUT2 deficiency. *See* Fucosyltransferase 2 deficiency (FUT2 deficiency)
- FVIII-replacement therapy, 346
- FVL. *See* Factor V Leiden (FVL)
- FXN gene. *See* *Frataxin* gene (FXN gene)
- Fy glycoprotein, 417
- Fya, 417
- Fyb, 417
- G**
- G protein  $\alpha$ -subunit (GNAS), 227–228, 228*t*
- G protein-coupled receptors (GPCR), 227–228
- G6PD deficiency, 417
- GA. *See* Glutaric aciduria (GA)
- Gamma dysfibrinogenemia. *See* Congenital gamma-chain molecular defect
- Gamma-tubulin ring complex ( $\gamma$ TuRC), 39
- GAP domain-containing protein. *See* GTPase-activating domain-containing protein (GAP domain-containing protein)
- Gap junctions, 20
- Gas chromatography coupled to mass spectrometry (GC–MS), 48, 51
- GAS1 genes, 92
- Gastrectomy, 84–85
- Gastrointestinal side effects (GI side effects), 226
- Gastrointestinal tract, 79–80
- GATA3 gene, 126
- GATA4 gene, 126
- Gaucher disease, 508
- GBA mutations. *See* Glucocerebrosidase mutations (GBA mutations)
- GCH1 mutations. *See* Dopa-responsive dystonia (DYT5)
- GCK. *See* Glucokinase (GCK)
- GC–MS. *See* Gas chromatography coupled to mass spectrometry (GC–MS)
- GEF. *See* GTP exchange factor (GEF)
- Gender assignment, 116–117
- Gene therapy, 98–101, 511–512, 511*f*
- achondroplasia, 100
- DMD, 99–100
- ex vivo and in vivo gene therapy, 512*f*
- hemophilia B, 100
- laminopathies, 101
- Leber congenital amaurosis, 100
- primary immunodeficiency diseases, 99
- sickle cell disease, 101
- Gene transcription, 23–29
- “Gene-specific” diagnoses, 535
- Gene(s), 3–7
- deletions in  $\beta$  thalassemia, 330
- ( $\epsilon\gamma\delta\beta$ )<sup>0</sup> thalassemia, 330
- restricted to  $\beta$ -globin gene, 330
- upstream deletions, 330
- discovery, 171
- editing, 516–519
- flow, 10
- gene-driven approaches, 12
- gene-environment interactions, 170–171
- implicated in Parkinson's disease, 464
- mutation, 171
- panels, 301
- silencing or suppression, 99
- switch during fetal to adult transition, 328
- transcription, 475
- Generation–sequencing techniques, 342–343
- Genetic heterogeneity, 92–96
- in CKD, 439
- splitting, 92–94
- ciliopathies, 95–96
- filaminopathies, 96



- Genetic heterogeneity (*Continued*)  
  holoprosencephaly, 92–93  
  laminopathies, 94–95  
  RASopathies, 94  
  SHFM, 93–94  
  syndromes with genetic connections, 94–96
- Genetic risk scores (GRSs), 184
- Genetic studies, 354  
  of hypertension, 184  
  twin, 355–356, 356*t*
- Genetic susceptibility, 487  
  to severe influenza, 411
- Genetic testing, 116, 161–164, 298, 437, 476, 532  
  for cancer, 431–434  
  genetic basis of  
    inheritable cardiomyopathies, 161–162  
    genes and encoded proteins associated with genetic cardiomyopathies, 162*t*  
  genetics of acquired heart failure, 161–164  
  indications for genetic testing in cardiomyopathies, 162–163  
  cardiac genetics multidisciplinary team, 163  
  diagnostic confirmation and prognostication, 162–163  
  genetic testing in heart failure, 163–164  
  genetic testing techniques in heart failure, 163  
  predictive testing of family members, 163  
  methods, 301–302  
    in heart failure, 163  
    limitations, 302–304  
  in nephrology, 439  
  role, 155  
  by Sanger sequencing, 147
- Genetic(s), 3–7, 253, 370–373, 462*t*  
  analysis of stroke, 192  
  of ataxia, 475  
  autosomal dominant inheritance, 461–463  
  basis of individual drug-response variation, 523–525  
  cardiomyopathies, 154, 156  
  code, 4–6  
  counseling, 86, 331, 441  
    for inherited cancer susceptibility, 430–431  
  diagnosis, 75, 77  
  diseases, 205  
  disorders, 90, 98, 427  
  distance, 9  
  drift, 10  
  of dystonia, 465  
  etiology, 89  
  factors, 184, 192–193, 201, 205–207  
    in pharmacodynamics, 524–525  
    in pharmacokinetics, 524  
  generalized epilepsies. *See* Idiopathic generalized epilepsy  
  and genomic studies in bronchial asthma, 354–363  
  candidate gene studies, 357–359  
  genetic linkage, 356–357
- GWAS, 359–362  
  NGS, 362–363  
  segregation analysis, 354–355  
  twin genetic studies, 355–356, 356*t*
- imprinting and cancer, 426–427
- linkage, 356–357, 358*t*, 360*t*  
  chromosome 5q, 357  
  chromosome 6p, 357  
  chromosome 11, 357  
  chromosome 12q, 357
- markers, 370
- medicine, 98
- obesity, 74, 81  
  POMC, 81
- overlap of monogenic and essential hypertension, 183–184
- polymorphisms, 189, 524
- predisposition, 393
- profiling, 501
- revolution in medicine, 535–536
- risk factors, 205
- of stroke, 189, 191–207
- syndromes, 90
- variation/variants, 170–171, 183–184 and CKD, 440
- Genital ambiguity, 110–111
- Genome, 369–370  
  genome-based microbial infection diagnostic kit, 394  
  genome-driven diagnoses, 535  
  genome-wide approach, 359  
  screen approach, 357  
  sequencing, 409, 536
- Genome editing, 516–519  
  CRISPR/Cas9-mediated DNA cleavage, 518*f*
- Genome-wide association studies (GWAS), 122, 171, 192, 248–249, 272, 293–294, 314, 359–362, 361*t*, 371, 394, 404, 423, 426, 438, 480, 490–491  
  metaanalysis, 371*f*  
  results from, 490–492
- Genome-wide associations, 361–362
- Genome-wide linkage analysis, 412  
  study, 196
- Genome-wide screening, 356
- Genomics, 3–7, 68–69, 370–373, 438  
  disorders, 91  
  DNA testing, 442  
  medicine, 14  
  ethical issues in, 532  
  revolution in medicine, 535–536  
  techniques, 442
- Genotype-first approach, 91
- Genotype–phenotype correlation, 91, 333–334
- Genotyping, 303
- Germ line, 511–512
- GH. *See* Growth hormone (GH)
- GI side effects. *See* Gastrointestinal side effects (GI side effects)
- Giant platelets, 349
- Glanzmann's thrombasthenia, 349
- Gliosis, 451–452
- Globin  
  chains, 327  
  gene clusters, 327  
  polypeptide, 327
- Globular monomeric actin molecules (G-actin), 21
- Glucocerebrosidase mutations (GBA mutations), 463
- Glucocorticoid remediable aldosteronism (GRA), 171, 181
- Glucocorticoids, 79
- Glucokinase (GCK), 283
- Glucose transporters (GLUTs), 268  
  GLUT-1, 268  
    deficiency syndrome, 301*b*  
  GLUT-2, 268  
  GLUT-3, 268  
  GLUT-4, 268, 269*f*
- Glucose utilization, 267–268
- Glucose-6-phosphate dehydrogenase deficiency, 524
- Glutaric aciduria (GA), 48
- Glutaryl-CoA dehydrogenase, 49
- GLUTs. *See* Glucose transporters (GLUTs)
- Gly-Pro-X motifs, 202
- Glycemic homeostasis  
  cellular glucose transport, 268  
  in fasting and postprandial state, 268*f*  
  glucose utilization, 267–268  
  insulin and insulin receptor, 268–270  
  molecular basis of, 267–270
- Glycemic regulation, disorders of, 270–274  
  autoimmunity and T1DM, 272–273  
  environmental factors, 273  
  HLA and T1DM, 272  
  molecular mechanisms in type 1 diabetes mellitus, 270–272  
  other genes or gene regions, 272  
  T2DM, 273–274
- Glycerol, 17
- Glycolysis, 43, 417
- Glycosylated hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>), 276
- Glycosylated proteins, 505–506
- Glycosylphosphatidyl-inositol (GPI), 18, 482
- GNA11, 225
- GNAS. *See* G protein  $\alpha$ -subunit (GNAS)
- GnRH. *See* Gonadotropin-releasing hormone (GnRH)
- Golgi apparatus, 36–37
- Golgi complex, 36–37
- Gonadal cancer risk, 117
- Gonadal differentiation, 109*f*, 110
- Gonadotropin-releasing hormone (GnRH), 219–220
- Gordon's syndrome, 171–175  
  case report, 171  
  diagnosis, 175  
  genetics, 172–173  
  management, 175  
  monogenic forms of hypertension, 172*t*  
  pathophysiology, 173–175
- Gorlin syndrome, 424
- GPCR. *See* G protein–coupled receptors (GPCR)
- GPI. *See* Glycosylphosphatidyl-inositol (GPI)



- GRA. *See* Glucocorticoid remediable aldosteronism (GRA)
- Graft-versus-host disease (GVHD), 315
- Grand mal, 290
- gRNA. *See* Guide RNA (gRNA)
- Growth hormone (GH), 217
- Growth-promoting genes, 425
- GRSs. *See* Genetic risk scores (GRSs)
- GTP. *See* Guanosine 5'-triphosphate (GTP)
- GTP exchange factor (GEF), 33–34
- GTPase eEF1A, 34
- GTPase eEF2, 34
- GTPase protein family members and biological processes, 30, 31†
- GTPase-activating domain-containing protein (GAP domain-containing protein), 42
- Guanine (G), 4
- Guanosine 5'-triphosphate (GTP), 19
- Guide RNA (gRNA), 517
- GVHD. *See* Graft-versus-host disease (GVHD)
- GWAS. *See* Genome-wide association studies (GWAS)
- γ-globin variants, 335–336
- H**
- H2A histone family, and member Y (H2AFY), 457
- H2BK120, ubiquitination of, 129
- H3K4. *See* Histone 3 lysine 4 (H3K4)
- H5N1, 411
- HAND1, 126
- Hansen's disease. *See* Leprosy
- Hantavirus cardiopulmonary syndrome (HCPS), 411
- Hantavirus infection, 411–412
- genetic susceptibility to severe influenza, 411
- immune-related gene expression variability and severity, 411
- immunity-related gene polymorphisms and severity, 411
- influenza-associated encephalopathy, 412
- Haploid, 6
- Haplotype-specific expression, 316–317
- Hapten (prohaptin) mechanism, 313
- Hb disorders. *See* Hemoglobin disorders (Hb disorders)
- HbA<sub>1c</sub>. *See* Glycosylated hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>)
- HbA1C correlation, 276
- HbC. *See* Hemoglobin C (HbC)
- HbE. *See* Hemoglobin E (HbE)
- HBeAg seroconversion, 406
- HBGAs. *See* Histo-blood group antigens (HBGAs)
- HbH disease, 332–333
- HbH disease. *See* Hemoglobin H disease (HbH disease)
- HbM. *See* Hemoglobin M (HbM)
- HbAg clearance, 406
- HBV-related hepatocarcinogenesis, 406
- HBV-related hepatocellular carcinoma, 406
- HCC. *See* Hepatocellular carcinoma (HCC)
- hCG. *See* Human chorionic gonadotropin (hCG)
- HCM. *See* Hypertrophic cardiomyopathy (HCM)
- HCPS. *See* Hantavirus cardiopulmonary syndrome (HCPS)
- HCV infections. *See* Hepatitis C virus infections (HCV infections)
- HD. *See* Huntington's disease (HD)
- HDAC inhibitors. *See* Histone deacetylase inhibitors (HDAC inhibitors)
- HDL. *See* High-density lipoprotein (HDL)
- HDL-C. *See* High-density lipoprotein-cholesterol (HDL-C)
- Head and neck PGL (HN PGL), 236
- Heart development, 121–122
- Heart failure, 153. *See also* Cardiovascular developmental disorder
- causes of, 154†
- clinical assessment, 155–160
- cardiac biopsy, 160
- electrocardiography, 156–157
- exercise testing, 160
- family history, 155
- imaging studies, 157–160
- laboratory studies, 155
- personal history, 155
- physical examination, 155
- epidemiology, 153
- etiology of, 153–154
- genetic testing, 161–164
- genetics of acquired, 161
- management, 164–166
- exercise, 165–166
- nonsurgical device therapies, 165
- pharmacological therapies, 164–165
- risk stratification for sudden death, 165
- HeFH. *See* Heterozygous familial hypercholesterolemia (HeFH)
- Hematopoietic stem cell (HSC), 99
- Hematopoietic stem cell transplantation (HSCT), 99, 381, 416
- Heme
- biosynthesis, 336
- defects of erythroid, 336
- systemic regulation of, 418
- Hemoglobin (Hb). *See also* Blood alterations, 417–418
- electrophoresis, 332–333
- high-oxygen affinity, 336
- switch, 328
- Hemoglobin C (HbC), 335
- Hemoglobin disorders (Hb disorders), 327.
- See also* Bleeding disorder
- α thalassemia, 332–334
- β thalassemia
- gene deletions in, 330
- laboratory diagnosis, 331
- molecular causes, 330–331
- prevention, 331–332
- classification and genetics of thalassemias, 328–330
- Hb molecule, 327–328, 336
- β-globin gene expression and control, 328
- characteristics of α-globin and β-globin gene loci, 327–328
- gene switch during fetal to adult transition, 328
- globin gene clusters, 327
- organization of human globin gene clusters on chromosomes 11, 328†
- qualitative defects, 334–336
- Hemoglobin E (HbE), 335
- Hemoglobin H disease (HbH disease), 332–333, 335
- Hemoglobin M (HbM), 335–336
- Hemoglobinopathies, 417–418
- Hemophilia A, 345–346
- Hemophilia B, 100, 348
- Hemophilia B Leiden phenotype, 342, 348
- Hemophilus*, 362
- Hemorrhagic stroke, 189
- Hemostasis, 341, 346, 350
- regulatory proteins of, 341
- Hepatic nuclear factor (HNF), 283
- HNF1A, 283
- HNF4A, 283
- Hnf4a* gene, 379–380
- Hepatitis B recombinant vaccine, 507
- Hepatitis C virus infections (HCV infections), 403
- host genetic determinants in, 407
- Hepatocellular carcinoma (HCC), 405
- Hepatosplenomegaly, 331
- HER. *See* Human epidermal growth factor (HER)
- Hereditary ataxias, 475
- Hereditary factors, 3–7
- structure and organization of nucleic acids, 4–7
- Hereditary hemorrhagic telangiectasia, 348
- Hereditary spherocytosis, 417
- hERG gene. *See* Human ether-a-go-go related gene (hERG gene)
- Heritability (h<sup>2</sup>), 353
- Herpes simplex virus (HSV), 403
- hESCs. *See* Human embryos stem cells (hESCs)
- Heterogeneity in CKD, 439
- Heterogeneous nuclear proteins (hnRNP), 29
- Heterogenous pathogenic variants, 348
- Heteroplasmy, 11, 298
- Heterozygotes, 10, 330, 335
- Heterozygous familial hypercholesterolemia (HeFH), 254†
- HGP. *See* Human Genome Project (HGP)
- HIES. *See* Hyper-IgE syndrome (HIES)
- HIF. *See* Hypoxia-inducible factors (HIF)
- High-density lipoprotein (HDL), 246–247
- High-density lipoprotein-cholesterol (HDL-C), 247
- High-grade squamous intraepithelial lesions (HSIL), 408
- High-oxygen affinity Hbs, 336
- High-performance liquid chromatography (HPLC), 331
- High-resolution-MS techniques, 51–52
- High-throughput sequencing (HTS), 57, 487, 492
- Histidine-encoding allele, 409

- Histo-blood group antigens (HBGAs), 419  
 Histology agnostic treatment, 499–500  
 Histone 3 lysine 4 (H3K4), 129  
 Histone deacetylase inhibitors (HDAC inhibitors), 99, 101–102  
   ADPKD, 101–102  
   KS, 101  
 Histone modification, 98*t*, 129  
 Histopathological, 160  
 HIV. *See* Human immunodeficiency virus (HIV)  
 HLA. *See* Human leukocyte antigen (HLA)  
 HLA-A2. *See* Human leukocyte antigen-A2 (HLA-A2)  
 HLA-C expression. *See* Human leukocyte antigen-C expression (HLA-C expression)  
 HLA-DP expression. *See* Human leukocyte antigen-DP expression (HLA-DP expression)  
 HLHS. *See* Hypoplastic left heart syndrome (HLHS)  
 HMGCL deficiency. *See* 3-Hydroxymethyl glutaryl-CoA lyase deficiency (HMGCL deficiency)  
 HN PGL. *See* Head and neck PGL (HN PGL)  
 HNF. *See* Hepatic nuclear factor (HNF)  
 hnRNP. *See* Heterogeneous nuclear proteins (hnRNP)  
 Holmes–Gang syndrome, 90–91  
 Holoprosencephaly, 92–94  
   genetic causes of human, 93*t*  
 Holt–Oram syndrome, 125  
 Homeostasis of molecules, 49  
 Homologous recombination (HR), 499  
 Homozygotes, 330, 335  
 Homozygous sickle-cell patients, 336  
 Hormonal factors, 205  
 Hormone deficiency, 219*t*  
 Host genetic determinants of HIV infection, 404  
 Host receptors, 419  
 HPA. *See* Human platelet antigen (HPA)  
 HPG axis.  
   *See* Hypothalamic–pituitary–gonadal axis (HPG axis)  
 HPLC. *See* High-performance liquid chromatography (HPLC)  
 HPT. *See* Hyperparathyroidism (HPT)  
 HPT-jaw tumor syndrome (HPT-JT syndrome), 223  
 HPV. *See* Human papillomavirus (HPV)  
 HR. *See* Homologous recombination (HR)  
 HSC. *See* Hematopoietic stem cell (HSC)  
 HSCT. *See* Hematopoietic stem cell transplantation (HSCT)  
 HSIL. *See* High-grade squamous intraepithelial lesions (HSIL)  
 HSV. *See* Herpes simplex virus (HSV)  
 HtrA serine peptidase 1 gene (*HTRA1*), 190  
 HTS. *See* High-throughput sequencing (HTS)  
 Human chorionic gonadotropin (hCG), 220  
 Human embryogenesis, 96–97  
 Human embryos stem cells (hESCs), 509  
 Human epidermal growth factor (HER), 526  
   Her2, 49–50, 514, 526  
 Human ether-a-go-go related gene (hERG gene), 140  
 Human Genome Project (HGP), 7, 522–523  
 Human genome variation and disease, 7–10  
   measuring genetic and genomic variation, 8–9  
 Human genomics for socioeconomic development, 14  
 Human health, 4  
 Human immunodeficiency virus (HIV), 393, 403, 525  
   HIV-1, 403  
   mother-to-child transmission, 404–405  
   vaccines, 507–508  
 Human leukocyte antigen (HLA), 270, 272, 299–300, 310–312, 343, 377–378, 405, 442, 525, 527  
   alleles, 532  
   and blood transfusion, 322–323  
   class I molecules, 310, 407  
   class II molecules, 311, 407  
   and disease, 312–314, 313*t*  
   and drug-induced hypersensitivities, 313  
   epistatic interaction of MHC genes, 314  
   expression, 314–317  
     correlation in autoimmune diseases, 315  
     HLA-C expression and disease development, 314–315  
     HLA-DP expression and increased risk of chronic HBV infection, 315  
     low *vs.* high expression mismatches in transplantation, 315  
     mechanisms underlying allele-specific HLA expression, 316–317  
   gene variants, 406–407  
     and early HBeAg seroconversion, 406  
     efficacy of interferon alfa and NAs treatment, 406–407  
   HBV-related hepatocellular carcinoma, 406  
   host genetic determinants in hepatitis C virus infection, 407  
   response to hepatitis B virus vaccine, 406  
   risk of developing liver cirrhosis, 406  
   and spontaneous HBsAg clearance, 406  
   genes, 372, 408  
   genotypes alter mother-to-child transmission, 405  
   genotyping, 299–300  
   haplotypes, 395  
   HLA-B27, 407  
   HLA-DQ locus rs9275572, 406–407  
   HLA-DR15 haplotype, 316–317  
   HLA–antibody-detection techniques, 319–322  
     anti–human leukocyte antigen antibodies relevance, 320  
     non–human leukocyte antigen antibodies role, 320–321  
     preventive measures, 321–322  
     and organ transplantation, 317–319  
   allorecognition, 317–318  
   AMR, 318, 318*f*  
   antibody-dependent cell-mediated cytotoxicity, 319  
   cellular rejection, 319  
   classification of rejection, 318  
   complement activation, 318–319  
   direct activation of endothelium, 319  
   profiling, 394  
   testing in clinical practice  
     abacavir use in HIV infection, 527  
     for prediction Stevens–Johnson syndrome, 527  
   typing for renal transplantation, 442  
 Human leukocyte antigen-A2 (HLA-A2), 309  
 Human leukocyte antigen-C expression (HLA-C expression), 314–315  
 Human leukocyte antigen-DP expression (HLA-DP expression), 315  
 Human papillomavirus (HPV), 408, 507, 507*f*  
 Human platelet antigen (HPA), 322  
 100kGP project, 69  
 Hunger hormone, 79–80  
 Huntington’s disease (HD), 372, 447, 461, 480  
   treatment, 480  
 Hydrolase substrates, 38  
 Hydrophobic heme pocket, 336  
 8-Hydroxy-2-deoxyguanosine, 458  
 18-Hydroxycortisol, 181–182  
 17-Hydroxylase deficiency, 234  
 3-Hydroxymethyl glutaryl-CoA lyase deficiency (HMGCL deficiency), 48  
 17-Hydroxyprogesterone (17-OH), 116  
 Hydroxyurea, 335  
 Hyper-IgE syndrome (HIES), 416  
 Hypercholesterolemia  
   encountered causes of primary, 252*t*  
   familial, 252, 254*t*  
   management of, 256–260  
   polygenic, 249–251  
 Hyperhomocysteinemia, 345  
 Hyperparathyroidism (HPT), 222  
   primary, 222–226  
 Hyperphagia, 77  
 Hyperpigmented skin lesions, 425  
 Hyperplastic synovium, 396–397  
 Hyperpolarizing factor, 169–170  
 Hyperstimulation, 390  
 Hypertension, 10, 439  
 Hypertriglyceridemia, 249–251, 253, 255–256, 260  
   familial, 251*t*  
   genetic disorders resulting in, 255–256  
   management of, 260  
   primary lipoprotein disorders leading to, 255*t*  
 Hypertrophic cardiomyopathy (HCM), 518  
 Hyphenated techniques, 51  
 Hypokalemia, 140–141  
 Hyponatremia, 300  
 Hypoplastic left heart syndrome (HLHS), 121  
 Hypospadias, 220  
 Hypothalamic–pituitary–gonadal axis (HPG axis), 220  
 “Hypothesis-free” investigation, 57

- Hypothesis-independent genome-wide approach, 356
- Hypoxia-inducible factors (HIF), 238  
HIF-1 $\alpha$  and HIF-2 $\alpha$ , 238
- I**
- IA. *See* Invasive aspergillosis (IA)
- IBD. *See* Isobutyrylglycinuria (IBD)
- IBDs. *See* Inflammatory bowel disease (IBDs)
- ICD.  
*See* Implantable cardioverter–defibrillator (ICD)
- ICF. *See* Immunodeficiency, centromeric region instability, and facial anomalies syndrome (ICF)
- ID. *See* Intellectual disability (ID)
- Idiopathic generalized epilepsy, 291–293
- Idiopathic hyperplasia (IHA), 230
- IFN. *See* Interferon (IFN)
- IFN regulatory factor-1 (IRF-1), 407
- IFs. *See* Incidental findings (IFs)
- IgE. *See* Immunoglobulin E (IgE)
- IGS. *See* Internal guide sequence (IGS)
- IGT. *See* Impaired glucose tolerance (IGT)
- IHA. *See* Idiopathic hyperplasia (IHA)
- IL. *See* Interleukin (IL)
- IL-2 receptor complex locus (*IL2RA*), 272
- ILAE. *See* International League Against Epilepsy (ILAE)
- Illumina, 61
- Imaging  
studies, 157–160  
CMR imaging, 159–160  
echocardiography, 158–159  
imaging modalities, 160  
plain chest radiography, 157  
techniques, 456
- Immune tolerance induction therapy (ITI therapy), 346
- Immune/immunity  
immune-related gene expression variability, 411  
immunity-related gene polymorphisms, 411  
response, 418  
risk profile, 392–393
- Immunity-related GTPase gene (IRGM gene), 371–372
- Immunobiological functions, 311
- ImmunoChip, 372
- Immunodeficiency, centromeric region instability, and facial anomalies syndrome (ICF), 518
- Immunogenetics  
and cancer, 426  
and microbial infection, 407  
genes involved in adaptive immunity, 407  
genes involved in innate immunity, 407  
genes involved in T-cell regulation and function, 407
- Immunoglobulin E (IgE), 354
- Immunoglobulin superfamily, 21
- Immunological network, complexity of, 355
- Immunological platelet refractoriness, 322
- Immunosenescence, 393
- Immunostaining, 452–453
- Immunostimulatory oligonucleotides, 516
- Impaired glucose tolerance (IGT), 276
- Implantable cardioverter–defibrillator (ICD), 147, 165  
therapy, 165
- Imprinting, 426
- In silico panel (ISP), 65
- Inborn errors of metabolism, 47–49  
production of succinyl CoA from propionyl CoA, 48f
- Incidental findings (IFs), 62
- Inclusion body  $\beta$  thalassemia, 330
- Incontinentia pigmenti (IP), 380
- Indirect pathway, 317–318
- Individual drug-response variation, 523–525  
genetic factors  
in pharmacodynamics, 524–525  
in pharmacokinetics, 524
- Individualized medicine. *See* Precision medicine
- Individualizing drug therapy  
HLA testing in clinical practice, 527  
testing for CYP2C19 activity prior to use of clopidogrel, 528  
trastuzumab and ERBB2 genotype, 526  
warfarin use as anticoagulant, 526–527
- Individually rare genetic variants, 501
- Induced pluripotent stem cells (iPSCs), 509, 518  
medical usage of, 510f  
Yamanaka model of generation, 509f
- Inflammaging, 392–393  
molecular profiling of, 393
- Inflammatory arthritis, subgroups of, 394–395
- Inflammatory bowel disease (IBDs), 367, 369f, 394  
adaptive immune system, 379  
autophagy, 379  
causative genome variants, 378  
clinical implications and translation, 380–381  
clinical manifestation of, 369  
complex clinical predisposition with complex complications, 367–374  
CD with UC, 368f  
epidemiology, 369  
epigenome, 373  
genetics and genomics, 370–373  
genome *vs.* environment, 369–370  
metabolomics, 374  
proteomics, 374  
transcriptome, 373–374  
functional implications, 378  
genomic portrait of, 367  
IBD-associated biomarkers, 367, 374  
IBD-associated SNP markers, 372  
identification of NOD2 gene, 374  
inheritance of, 369–370  
MHC, 377–378  
mucosal barrier function, 379–380  
NOD2 and innate immunity, 375–377  
paradoxes with diseases, 380  
parallels with diseases, 380
- Inflammatory/inflammation, 389–390  
acute, 389  
chronic, 389–390  
control, 363  
disease, 380  
mediators, 390  
molecular biology of inflammatory processes, 400
- Influenza-associated encephalopathy, 412
- Inheritance of faulty oncogenes, 425
- Inherited arrhythmia syndromes, 133
- Inherited cancer, 423  
susceptibility, 424, 426  
genetic counseling for, 430–431
- Inherited kidney diseases, 437, 439
- Inherited macrothrombocytopenia, 349
- Inherited methemoglobinemia, 335–336
- Inherited monogenic diabetes mellitus, 283–284  
malformation syndromes with diabetes mellitus, 284  
mitochondrial diabetes mellitus, 284  
MODY, 283  
NDM, 283  
syndromes of inherited insulin resistance, 284
- Inherited susceptibility to leukemia, 427–430
- Initiation of protein synthesis, 34
- Innate immune/immunity, 407  
genes, 419  
genetic associations with HIV-1, 405  
response and cellular injury, 419–420  
acquired immunity, 420  
system, 390
- Inositol 1,4,5-trisphosphate (IP3), 19, 479
- INR. *See* International normalized ratio (INR)
- INS. *See* Insulin (INS)
- Institutional review boards (IRB), 532
- Insulin (INS), 268–270, 272, 285  
receptor, 268–270  
synthesis and secretion, 269f
- Insulinitis, 272–273
- Integral membrane protein 2B gene (*ITM2B* gene), 190
- Integral proteins, 18
- Integrated genomic and molecular medicine, 541, 542f. *See also* Personalized medicine (PM)  
evidence-based, precision, and personalized medicine, 537–539, 538f  
genetic, genomic, and molecular revolutions in medicine, 535–536  
stratified medicine, 539–541
- Integrins, 21
- Intellectual disability (ID), 489
- Interferon (IFN), 406–407  
alfa, 406–407
- Interleukin (IL), 272, 357, 407–408  
IL-1 $\beta$ , 203–204  
IL-8, 390  
IL-18, 438  
IL-22, 368–369  
IL-23R gene, 371–372, 379
- Intermediate filaments, 22

- Internal guide sequence (IGS), 514  
 International League Against Epilepsy (ILAE), 290  
 International normalized ratio (INR), 524  
 Intertumor heterogeneity, 501  
 Intervening sequences (IVSs), 327  
 Intracellular antiviral host factor affecting HIV disease, 405  
 Intracellular bacteria, host factors and genetic susceptibility to, 412–415  
   *Chlamydia pneumoniae*, 414  
   *Chlamydia trachomatis*, 413–414  
   *Coxiella burnetii*, 414–415  
   *Mycobacterium leprae*, 413  
   *Mycobacterium tuberculosis*, 412–413  
   *Mycoplasma pneumoniae*, 414  
   *Tropheryma whippelii*, 415  
 Intracellular membranes, 17  
 Intracerebral hemorrhage, 190  
   by candidate gene association studies, 201*t*  
   molecular genetics of, 201–202  
     apolipoprotein E gene, 201–202  
     collagen, type IV, alpha 1 gene, 202  
   molecular pathophysiology of, 199–201  
 Intracranial aneurysm  
   loci for, 207  
   molecular genetics of, 205–207  
   molecular pathophysiology of, 202–204  
 Intra-neuronal nuclear inclusions, 480  
 Intranuclear inclusions, 475  
 Intraperitoneal routes (IP routes), 375  
 Intratumor heterogeneity, 501  
 Intrauterine growth restriction (IUGR), 115  
 Introns, 4–6  
 Invasive aspergillosis (IA), 416  
 IP. *See* Incontinentia pigmenti (IP)  
 IP routes. *See* Intraperitoneal routes (IP routes)  
 IP3. *See* Inositol 1,4,5-trisphosphate (IP3)  
 iPSCs. *See* Induced pluripotent stem cells (iPSCs)  
 IRB. *See* Institutional review boards (IRB)  
 IRF-1. *See* IFN regulatory factor-1 (IRF-1)  
 IRGM gene. *See* Immunity-related GTPase gene (IRGM gene)  
 Iron chelation therapy, 328–329  
 Ischemic cardiomyopathy, 161  
 Ischemic stroke, 191  
   brain injury after, 193*f*  
   etiology, 193*f*  
   molecular genetics of, 194–199  
     arachidonate 5-lipoxygenase-activating protein gene, 197–198  
     CDKN2B antisense RNA 1 gene, 198  
     linkage analyses, 195*t*  
     Ninjurin 2 gene, 198–199  
     paired-like homeodomain 2 gene, 199  
     phosphodiesterase 4D, cAMP-specific gene, 196–197  
     zinc finger homeobox 3 gene, 199  
   molecular pathophysiology, 192–194  
   pathophysiology of brain injury after, 194*f*  
 Isobutyrylglycinuria (IBD), 49  
 Isolated congenital heart disease causing by single gene, 126  
 ISP. *See* In silico panel (ISP)  
 ITI therapy. *See* Immune tolerance induction therapy (ITI therapy)  
 ITM2B gene. *See* Integral membrane protein 2B gene (ITM2B gene)  
 IUGR. *See* Intrauterine growth restriction (IUGR)  
 IUGRhypopspadias syndrome, 115  
 IVSs. *See* Intervening sequences (IVSs)
- J**  
 Jervell and Lange-Nielsen syndrome, 141, 143  
 JMJD1C gene, 129  
 Joubert syndrome, 102  
 Juberg–Marsidi syndrome, 90–91  
 Juvenile myelomonocytic leukemia (JMML), 429  
 Juvenile Parkinson's disease, 39
- K**  
 K- $\alpha$ 1 tubulin, 320–321  
 K<sup>+</sup> voltage-gated channel subfamily J member 2 (KCNJ2), 141, 362–363  
 Kabuki syndrome (KS), 97, 98*t*, 101  
 KAL1. *See* Anosmin 1  
 Kallmann syndrome, 220  
 Karyotypic abnormalities, 296, 296*f*  
 Karyotyping, 302–303  
 KBG syndrome, 96–97, 98*t*  
 KCNE1 gene, 139, 141  
 KCNE2 gene, 140  
 KCNH2 gene, 140  
 KCNJ2. *See* K<sup>+</sup> voltage-gated channel subfamily J member 2 (KCNJ2)  
 KCNQ1 gene, 139, 141–142  
 KDM6A/UTX genes, 101  
 KDM7A gene, 129  
 Keratins, 22  
 Ketonuria, 275  
 Kidney transplantation, molecular basis of  
   donors selection, 442  
   HLA typing for renal transplantation, 442  
   predicting long-term kidney allograft function, 442  
 Killer immunoglobulin-like receptors (KIR), 314  
 Kinetochore microtubules, 41  
 KIR. *See* Killer immunoglobulin-like receptors (KIR)  
 Kisspeptin 1 (KISS1), 221  
 Kisspeptin receptor (KISS1R), 221  
 KLHL3 gene, 173–175  
 Klinefelter syndrome, 241  
 KMT2D/MLL2 genes, 101  
 KS. *See* Kabuki syndrome (KS)  
 Kynamro, 513
- L**  
 L chain of ferritin (*FtL*), 482  
 L858R. *See* Leucine with arginine at codon 858 in exon 21 (L858R)  
 Lacunar infarction, 192–193  
 LADA. *See* Latent autoimmune diabetes in adults (LADA)  
 LAM therapy. *See* Lamivudine therapy (LAM therapy)  
 Lamin A/C gene (LMNA gene), 161–162, 284  
 Laminopathies, 94–95, 101  
 Lamivudine therapy (LAM therapy), 406–407  
 Landau Kieffner syndrome, 294  
 Large curated panels (LCPs), 61–62  
 Larotrectinib, 499–500  
 Late endosomes, 37  
 Latent autoimmune diabetes in adults (LADA), 270  
 LBD. *See* Lewy body dementia (LBD)  
 LC. *See* Liver cirrhosis (LC)  
 LCAT. *See* Lecithin–cholesterol acyltransferase (LCAT)  
 LC–MS. *See* Liquid chromatography coupled to mass spectrometry (LC–MS)  
 LCPs. *See* Large curated panels (LCPs)  
 LCSD. *See* Left cardiac sympathetic denervation (LCSD)  
 LDL. *See* Low-density lipoprotein (LDL)  
 LDL-C. *See* Low-density lipoprotein cholesterol (LDL-C)  
 LDL-receptor-related proteins 5 and 6 (*LRP5* and *LRP6*), 113  
 LDLR. *See* Low-density lipoprotein receptor (LDLR)  
 Lecithin–cholesterol acyltransferase (LCAT), 246–247  
 Left cardiac sympathetic denervation (LCSD), 134, 148  
 Legius syndrome, 429  
 Leigh syndrome, 299  
 LEP. *See* Leptin (LEP)  
 LEPR. *See* Leptin receptor (LEPR)  
 Leprechaunism, 284  
 Lepromatous pole, 413  
 Leprosy, 413  
 Leptin (LEP), 75, 83, 221  
 Leptin receptor (LEPR), 77*b*  
   deficiency, 77–78  
   gene, 75, 78, 221  
 “Leptin–melanocortin” pathway, 80, 82–83, 276–277, 277*f*  
 Leucine with arginine at codon 858 in exon 21 (L858R), 498–499  
 Leucine-rich, immunoglobulin-like domain (LRIG3), 161  
 Leucine-rich repeat kinase 2 (LRRK2), 379, 462–463  
 Leukemia, inherited susceptibility to, 427–430  
 Leukotriene receptor antagonists, 363  
 Levodopa, 464  
 Lewy bodies, 451–452  
 Lewy body dementia (LBD), 452  
 LH. *See* Luteinizing hormone (LH)  
 LHX8 genes, 113  
 Libman–Sacks endocarditis, 414–415  
 Liddle's syndrome, 175–178  
   case report, 175  
   diagnosis, 177–178  
   genetics, 176



- management, 178  
 pathophysiology, 176–177
- Life-threatening disorders, 527
- Ligand(s), 18–19  
 ligand–receptor interaction, 18–19
- LIM domain kinase 1 gene (*LIMK1*), 205
- Linkage  
 analysis, 190, 192, 356  
 studies in families, 293
- Lipid bilayer, 17–18
- Lipodystrophy with insulin resistance, 284
- Lipoprotein, 245, 246*t*  
 agents in development, 260  
 apheresis, 259–260  
 classes, 247*f*  
 FH, 252, 254*t*  
 characteristic clinical features of, 252–254, 253*t*  
 screening strategies for, 255  
 genetic disorders resulting in  
 hypertriglyceridemia, 255–256, 255*t*  
 management  
 of hypercholesterolemia, 256–260  
 of hypertriglyceridemia, 260  
 metabolism, 245–247  
 cholesterol in, 248*f*  
 diagnosing, 249  
 environmental and genetic factors  
 affecting, 247–249  
 polygenic hypercholesterolemia, 249–251  
 screening for lipoprotein disorders, 249  
 treatment of primary lipoprotein disorders, 256  
 type 3 hyperlipoproteinemia, 256
- Liquid chromatography coupled to mass spectrometry (LC–MS), 51  
 application of, 51
- Lissencephaly, 22
- Liver cirrhosis (LC), 405–406
- Living organisms, 4
- LLPCs. *See* Long-lived plasma cells (LLPCs)
- LMNA gene. *See* Lamin A/C gene (LMNA gene)
- Local neutralization of cytokines, 393
- LOF. *See* Loss-of-function (LOF)
- Lomitapide, 259–260
- Long QT syndrome (LQTS), 133, 137–141, 147–148. *See also* Short QT syndrome (SQTS)  
 Anderson–Tawil syndrome, 141  
 clinical risk assessment in, 147  
 diagnosis, 138–141  
 drug-induced long QT, 143  
 generation and propagation of arrhythmia in, 143–144  
 Jervell and Lange-Nielsen syndrome, 141  
 LQTS type 1, 139–140  
 LQTS type 2, 140–141  
 LQTS type 3, 141, 148  
 molecular risk stratification in, 141–143  
 targeted therapies in, 147–148  
 Timothy syndrome, 141
- Long-acting beta 2 agonists, 363
- Long-lived plasma cells (LLPCs), 318
- Long-read sequencing, 372
- Long-term kidney allograft function, predicting, 442
- Loss of action potential dome, 146–147
- Loss-of-function (LOF), 512–513  
 LoF-intolerant genes, 490–491  
 mutation, 66
- Low-density lipoprotein (LDL), 245–246
- Low-density lipoprotein cholesterol (LDL-C), 258*t*
- Low-density lipoprotein receptor (LDLR), 38
- Low-oxygen affinity Hbs, 336
- LQT7, 141
- LQTS. *See* Long QT syndrome (LQTS)
- LRP5 and LRP6. *See* LDL-receptor-related proteins 5 and 6 (LRP5 and LRP6)
- LRRK2. *See* Leucine-rich repeat kinase 2 (LRRK2)
- LRRK2/PARK8, 462–463
- LTA. *See* Lymphotoxin- $\alpha$  (LTA)
- Luminex platform, antibody screening on, 319–320
- Lung cancer, 523
- Luteinizing hormone (LH), 217, 220
- Lymphotoxin *a*, 377–378
- Lymphotoxin- $\alpha$  (LTA), 413
- Lynch syndrome, 425
- Lysosomal hydrolytic enzymes, 37
- Lysosomal storage diseases, 38
- Lysosomes, 37–38
- M**
- M cells, 143
- M2-PK. *See* Pyruvate kinase type M2 (M2-PK)
- MAC. *See* Membrane attack complex (MAC)
- Macrophage inflammatory protein (MIP), 405
- Macrophage-stimulating protein (MST1), 378
- Magnetic resonance angiography (MRA), 205
- Magnetic resonance imaging (MRI), 190, 450
- Magnetic resonance spectroscopy (MRS), 456
- MAIS. *See* Mild androgen insensitivity syndrome (MAIS)
- Major depression, 492
- Major histocompatibility complex (MHC), 309, 377–378, 405  
 with chromosomal location and gene map, 310*f*  
 epistatic interaction of MHC genes, 314  
 features, 311*b*
- Malaria, 417–419  
 immune response, 418  
 influence of erythrocyte and immune response gene polymorphisms in, 418*f*  
 membrane and enzymatic disorders of red blood cells, 417–418  
 hemoglobinopathies, 417–418  
 systemic regulation of heme, 418  
 vaccine, 419
- Malformation syndromes  
 with diabetes mellitus, 284  
 with transcriptional dysregulation, 98*t*
- MAM domain containing glycosylphosphatidylinositol anchor 1 (MDGA1), 362–363
- Manic depressive disorder, 487, 489, 492
- Mannose-6-phosphate, 37
- Mannose-binding lectin protein (MBL protein), 405  
 MBL2 gene, 405
- MAP. *See* Microtubule-associated proteins (MAP)
- MAPK. *See* Mitogen-activated protein kinase (MAPK)
- Maple syrup disease (MSD), 48–49
- Mass spectrometry (MS), 49
- Massively parallel sequencing, 57
- Maternal androgen-producing tumors, 114
- Matrix metalloproteinases (MMPs), 203–204
- Maturity onset diabetes of the young (MODY), 270, 283, 380
- May–Hegglin anomaly, 349
- MBL protein. *See* Mannose-binding lectin protein (MBL protein)
- MC4R. *See* Melanocortin 4 receptor (MC4R)
- MCM8 genes, 113
- MCM9 genes, 113
- MDGA1. *See* MAM domain containing glycosylphosphatidylinositol anchor 1 (MDGA1)
- MDP. *See* Muramyl dipeptide (MDP)
- Mean fluorescence intensity (MFI), 319
- Mechanistic target of rapamycin (mTOR), 19, 32
- Mediator, 28
- Medullary thyroid cancer (MTC), 223, 237, 239
- MEF2C. *See* Myocyte enhancer factor 2C (MEF2C)
- Megakaryocyte progenitors, 429–430
- Meiosis, 6, 7*f*
- MEK inhibitor, 499
- Melanesian elliptocytosis. *See* South Asian ovalocytosis
- Melanocortin 4 receptor (MC4R), 75, 77*b*, 81  
 gene mutations, 75, 77
- Melanoma, 497–498
- MELAS. *See* Myoclonic epilepsy with lactic acidosis (MELAS)
- Membrane attack complex (MAC), 309
- Membranoproliferative glomerulonephritis, 440
- MEMSA. *See* Myoclonic epilepsy, myopathy and sensory ataxia (MEMSA)
- MEN. *See* Multiple endocrine neoplasia (MEN)
- Mendelian diseases, 356, 394
- Mendelian genetics, 535–536
- Mendelian inheritance, 293
- Mendelian susceptibility to mycobacterial disease (MSMD), 412
- Mendelian traits, 533
- meQTL. *See* Methylation QTL (meQTL)
- MERRF. *See* Myoclonic epilepsy with ragged red fibers (MERRF)
- Messenger RNA (mRNA), 4–6, 12  
 quality control, 30–32  
 translation, 32–34
- Met-tRNA<sup>Met</sup>, 34
- Metabolic



- Metabolic (*Continued*)  
 diseases/disorder, 49–50, 158–159  
 fingerprinting technique, 51  
 heart diseases, 160  
 syndrome, 249–251, 278
- Metabolite(s), 47, 392  
 profiling, 374
- Metabolome, 47
- Metabolomics, 50, 374
- Methemoglobinemic hemoglobin variants, 335–336
- Methylation of CpG islands, 128
- Methylation QTL (meQTL), 311
- Methylmalonic acidemia (MMA), 48
- Methyltetrahydrofolate reductase mutation (MTHFR mutation), 341
- Mexiletine, 148
- MFI. *See* Mean fluorescence intensity (MFI)
- MHC. *See* Major histocompatibility complex (MHC)
- MHC class I chain-related molecule A (MICA), 320–321
- MHC class-I-related chain gene family (MIC gene family), 377–378
- MICA. *See* MHC class I chain-related molecule A (MICA)
- Mice models, 377
- Microarray/biochips, 523
- Microbial infections, 403  
 dengue viral infection, 410–411  
 genetic susceptibility  
   to enteral pathogens, 419–420  
   of humans to hantavirus infection, 411–412
- HLA gene variants, 405  
 and susceptibility and persistence of HBV infection, 406–407
- host factors and genetic susceptibility to intracellular bacteria, 412–415
- host genetics  
 of Epstein–Barr infection, 409–410  
 susceptibility and protection from fungal infections, 415–417  
 susceptibility to HPV infection and progression to cervical cancer, 408
- immunogenetics and microbial infection, 407
- malaria, 417–419
- understanding host genetic variation of susceptibility, 403–405  
 chemokine receptor genetic variants, 404–405  
 coreceptor and ligand variants in HIV disease progression, 404  
 HLA genotypes alter mother-to-child transmission, 405  
 host genetic determinants of HIV infection, 404  
 innate immunity genetic associations with HIV-1, 405  
 intracellular antiviral host factor affecting HIV disease, 405  
 rate of disease progression, 405
- Microbiota, 370
- Microfluidics digital PCR, 332
- Microna-196 (miR-196), 379
- MicroRNAs (miRNAs), 123–124, 126–128, 127*t*, 374, 381, 457  
 binding, 316  
 miRNA-182, 126
- Microsatellite instability (MSI), 500
- Microtubule-associated proteins (MAP), 22
- Microtubule-organizing centers (MTOC), 22, 39–42  
 and cell cycle, 39–41, 40*f*  
 primary cilium, 41–42
- Microtubules, 22
- Miglustat, 483
- Mild androgen insensitivity syndrome (MAIS), 112–113
- Mild hemophilia A, 342
- MINA gene, 129
- Mini genome model, 309–310
- MinK1, 141
- MIP. *See* Macrophage inflammatory protein (MIP)
- Mipomersen, 259–260, 513
- miR-196. *See* Microna-196 (miR-196)
- miRNAs. *See* MicroRNAs (miRNAs)
- Mismatch repair system (MMR system), 425, 500
- Missense mutation, 6
- Missing inheritance, 372
- Mitochondria(l), 42–44, 298  
 diabetes mellitus, 284  
 disorders, 298  
 DNA, 42–43  
 dysfunction, 392–393, 454, 475  
 epilepsies, 298–299  
   clinical phenotypes, 298  
   heteroplasmy, 298  
   syndromes, 298–299  
 genes, 42–43  
 genome, 8, 9*t*, 10–11, 11*f*  
 testing, 302
- Mitochondrial DNA (mtDNA), 8–9  
 mutations in, 11  
 replication and transcription, 11
- Mitochondrial membrane protein-associated neurodegeneration (MPAN), 482
- Mitochondrion, 42–43
- Mitogen-activated protein kinase (MAPK), 497–498  
 signaling pathway, 103
- Mitosis, 6, 7*f*, 41
- Mixed gonadal dysgenesis, 110, 116
- MLPA analysis. *See* Multiplex ligation-dependent probe amplification analysis (MLPA analysis)
- MMA. *See* Methylmalonic acidemia (MMA)
- MMPs. *See* Matrix metalloproteinases (MMPs)
- MMR system. *See* Mismatch repair system (MMR system)
- Modern genetic testing, 438
- Modern sequencing methods, 440
- MODS. *See* Multiple organ dysfunction syndrome (MODS)
- MODY. *See* Maturity onset diabetes of the young (MODY)
- Molecular basis of individual drug-response variation, 523–525
- Molecular diagnostic(s), 342  
 application in neurodegeneration, 456–458  
 testing, 342–343
- Molecular genetic(s), 333–334  
 diagnosis of CKD, 439  
 epilepsy, 289–291  
   evidence for genetic basis, 291–292  
   genes, 302*b*  
   genetic architecture, 292–298  
   genetic testing methods, 301–302  
   limitations to, 302–304  
   management, 289  
   mitochondrial, 298–299  
   molecular genetic testing strategies, 300–304  
   pharmacogenetics, 299–300  
   progressive myoclonic, 299  
 seizures, 289–290
- Molecular medicine, 91, 505
- Molecular pathology, 395  
 of neurodegenerative diseases, 453–455
- Molecular revolution in medicine, 535–536
- Molecular risk stratification  
 in CPVT, 145–146  
 in long QT syndrome, 141–143
- Molecular testing, 342–343, 403
- Molecular therapy  
 acute inflammatory diseases, 400  
 chronic inflammatory diseases, 400
- Molecularly targeted therapy limitations in cancer, 500–501
- Mono-ubiquitin, 39
- Monoclonal antibodies, 363
- Monogenic disorders, 74, 327
- Monogenic forms of systemic hypertension, 171–181  
 congenital adrenal hyperplasia, 178–180  
 diagnosis, 181  
 Gordon's syndrome, 171–175  
 Liddle's syndrome, 175–178
- Monogenic mutations, 372
- Monogenic obesity syndromes with intellectual disability, 81–83  
 BBS, 82–83  
 PWS, 81–82, 82*f*
- Monogenic traits, 533
- Monosomy 7, 427
- Monozygotic twins (MZ twins), 354, 369–370
- Moraxella, 362
- Mother centrioles (*M. centrioles*), 40
- Mother-to-child transmission (MTCT), 404–405
- Movement disorders, 461  
 ataxia, 472–479  
 DRPLA, 483  
 dystonia, 464–471  
 HD, 480  
 neurodegeneration with brain iron accumulation, 481–483  
 Niemann–Pick type C, 483  
 Parkinson's disease, 461–464  
 Wilson's disease, 480–481

- MPAN. *See* Mitochondrial membrane protein-associated neurodegeneration (MPAN)
- MRA. *See* Magnetic resonance angiography (MRA)
- MRI. *See* Magnetic resonance imaging (MRI)
- mRNA. *See* Messenger RNA (mRNA)
- MRS. *See* Magnetic resonance spectroscopy (MRS)
- MS. *See* Mass spectrometry (MS)
- MSD. *See* Maple syrup disease (MSD)
- MSI. *See* Microsatellite instability (MSI)
- MSI high (MSI-H), 500
- MSMD. *See* Mendelian susceptibility to mycobacterial disease (MSMD)
- MST1. *See* Macrophage-stimulating protein (MST1)
- MTC. *See* Medullary thyroid cancer (MTC)
- MTCT. *See* Mother-to-child transmission (MTCT)
- mtDNA. *See* Mitochondrial DNA (mtDNA)
- MTHFR C677T polymorphism, 122
- MTHFR mutation.  
     *See* Methyltetrahydrofolate reductase mutation (MTHFR mutation)
- MTOC. *See* Microtubule-organizing centers (TOC)
- mTOR. *See* Mechanistic target of rapamycin (mTOR)
- Mucin-2 (MUC2), 379–380
- Mucosal barrier function, 379–380
- Multibacillary leprosy. *See* Lepromatous pole
- Multicistronic RNAs, 11
- Multiple artificial nuclease systems, 517
- Multiple endocrine neoplasia (MEN), 425  
     clinical case, 222  
     clinical characteristics, 223*t*, 224*t*  
     MEN1, 224–225, 425  
     MEN2, 239, 425  
         MEN2A syndrome, 236, 239  
         MEN2B syndrome, 236, 239  
     MEN4, 223  
     molecular systems, 223–225  
     syndromes, 222–226
- Multiple organ dysfunction syndrome (MODS), 390
- Multiple resistance mechanisms, 499
- Multiplex ligation-dependent probe amplification analysis (MLPA analysis), 333, 334*f*
- Multisystem chronic inflammatory disease, 398
- Multisystemic manifestations, 48
- Multivariate mathematical models, 249
- Muramyl dipeptide (MDP), 375
- Muscle wasting, 397–398
- Mutation(s), 10, 329, 341  
     in  $\alpha$  gene, 332  
     in mtDNA, 11
- MYBPC3 gene, 518
- Mycobacteria, 376–377
- Mycobacterial diseases, 393–394
- Mycobacterium*  
     *M. leprae*, 380, 413  
     *M. tuberculosis*, 412–413
- Mycoplasma pneumoniae*, 414
- Myelofibrosis, 430
- Myeloproliferative disorders, 427
- Myocardial disease, 160
- Myocardial pathology, 156
- Myoclonic epilepsy, myopathy and sensory ataxia (MEMSA), 298
- Myoclonic epilepsy with lactic acidosis (MELAS), 284, 298–299
- Myoclonic epilepsy with ragged red fibers (MERRF), 298
- Myocyte enhancer factor 2C (MEF2C), 362–363
- Myosin chains, 21
- Myosin heavy chain (MYH7), 161–162
- Myosin-binding protein C (MYBPC3), 161–162
- MZ twins. *See* Monozygotic twins (MZ twins)
- N**
- N*-methyl-D-aspartate receptor (NMDAR), 492
- N*-terminal pro-BNP, 155
- $\text{Na}^+ - \text{Ca}^{2+}$  exchange pump (NCX), 145
- NADH. *See* Nicotinamide adenine dinucleotide (NADH)
- Naïve T cells, 407–408
- Narcolepsy, 312
- NARP. *See* Neuropathy, ataxia, and retinitis pigmentosa (NARP)
- National Health Service (NHS), 394, 541
- National High Blood Pressure Education Program, 169
- National Institute for Health and Clinical Excellence (NICE), 226, 249, 394
- National Institute of Health (NIH), 192, 225, 537–538
- National Institute of Health Research (NIHR-UK), 537–538
- Natriuretic peptides, 170
- Natural selection, 10
- NBD. *See* Nucleotide-binding domain (NBD)
- NCX. *See*  $\text{Na}^+ - \text{Ca}^{2+}$  exchange pump (NCX)
- NDM. *See* Neonatal diabetes mellitus (NDM)
- Neisseria gonorrhoeae*, 403
- Neonatal diabetes mellitus (NDM), 283
- Neurobiology, 448
- Neurocognitive dysfunction, 440
- Neurodegenerative/neurodegeneration, 447  
     aceruloplasminaemia, 482  
     beta-propeller protein-associated neurodegeneration, 482  
     with brain iron accumulation, 481–483  
     COASY protein-associated neurodegeneration, 482  
     diseases, 447  
     clinical case studies and molecular systems, 448–453  
     FAHN, 482  
     Kufor Rakeb, 483  
     MPAN, 482  
     neuroferritinopathy, 482  
     PKAN, 481  
     PLAN, 481
- Neuroferritinopathy, 482
- Neurofibromatosis type 1 (NF1), 426
- Neuronal calcium homeostasis, 475
- Neuropathy, ataxia, and retinitis pigmentosa (NARP), 298
- Neuropsychiatric disorders, molecular pathology in  
     CNVs, 487–489  
     cognitive function among CNVs carriers, 489  
     high-throughput sequencing studies, 492  
     pathway analysis, 493  
     penetrance of CNVs, 489–490  
     results from GWAS, 490–492
- Neurotransmitter receptors, 491–492
- Neurotrophic receptor tyrosine kinase fusions (NTRK fusions), 500
- Neutral mutations, 6
- Neutrophil-mediated tissue-injury theory, 389
- Next-generation sequencing (NGS), 57–59, 91, 123, 147, 345, 362–363, 394, 437, 461, 465  
     bioinformatics, 65–66  
     choice of test, 59–62  
     clinical interpretation of variants, 66–69  
     ethical considerations, 62–65  
     Sanger sequencing, 58  
     second-generation sequencing, 58–59  
     third-generation sequencing, 59
- NF1. *See* Neurofibromatosis type 1 (NF1)
- NGS. *See* Next-generation sequencing (NGS)
- NHS. *See* National Health Service (NHS)
- NICE. *See* National Institute of Clinical Excellence (NICE)
- Nicolaides–Baraitser syndrome, 98*t*
- Nicotinamide adenine dinucleotide (NADH), 43, 335–336
- Nicotinic acid, 260
- Niemann–Pick C1-like 1 receptors, 259
- Niemann–Pick type C, 483
- NIH. *See* National Institute of Health (NIH)
- NIHR-UK. *See* National Institute of Health Research (NIHR-UK)
- Ninjurin 2 gene (*NINJ2*), 195, 198–199
- Nitric oxide (NO), 18–19, 169–170
- Nitric oxide synthase 2, inducible (NOS2), 203–204
- Nitric oxide synthase adaptor protein (*NOS1AP*), 143
- Nivolumab, 500
- NKX2.5 gene, 122, 126
- Nlrp3 inflammasome, 392
- NMDAR. *See* *N*-methyl-D-aspartate receptor (NMDAR)
- NMR spectra. *See* Nuclear magnetic resonance spectra (NMR spectra)
- NO. *See* Nitric oxide (NO)
- NOBOX genes, 113
- NOD2. *See* Nucleotide binding oligomerization domain-2 (NOD2)
- NODAL genes, 92
- Noncardiac congenital anomalies in CHD, 128
- Noncardioselective  $\beta$ -blocker drug, 147–148

- Nonclassic CAH, 233  
 Noncoding regulatory genome in CHD, 126–128  
   and single-nucleotide polymorphisms, 127–128  
 Noncoding variants, 295–296  
 Nondeletional  $\beta$  thalassemia  
   abnormal posttranscriptional modification of mRNA, 329  
   molecular basis of, 329  
   mRNA synthesis, 329  
   mutants affecting mRNA processing, 329  
   mutations resulting in abnormal posttranscriptional modification of mRNA, 329  
 Nonhead and neck paragangliomas (Non-HN PGLs), 236  
 Non-human leukocyte antigen antibodies, 320–321  
 Nonhypothesis genomic methodologies, 370  
 Noninvasive diagnostic methods, 447  
 Nonpersonalized medication, 85  
 Nonsense mutation, 6  
 Non-small cell lung cancer (NSCLC), 497–498, 523, 540  
 Nonsteroidal antiinflammatory drugs, 400  
 Nonsynonymous variants, 362  
 Noonan syndrome (NS), 125, 425  
 NOS1AP. *See* Nitric oxide synthase adaptor protein (NOS1AP)  
 NOS2. *See* Nitric oxide synthase 2, inducible (NOS2)  
 NOTCH1 gene, 122, 126  
 NOTCH3 gene, 190  
 NR5A1 gene, 110–111  
 NRAMP1 gene polymorphisms, 413  
 NS. *See* Noonan syndrome (NS)  
 NSCLC. *See* Non-small cell lung cancer (NSCLC)  
 NTRK fusions. *See* Neurotrophic receptor tyrosine kinase fusions (NTRK fusions)  
 Nuclear envelope, 30–32  
 Nuclear genes, 7  
 Nuclear genome, 8, 9*t*  
 Nuclear imaging techniques, 160  
 Nuclear lamina, 23–29  
 Nuclear magnetic resonance spectra (NMR spectra), 49, 51–52, 456  
 Nucleic acids, 4–7, 535–536  
 Nucleolar organizing regions, 30  
 Nucleolus, 30  
 Nucleosome, 24–28  
 Nucleotide ATP, 42  
 Nucleotide binding oligomerization domain-2 (NOD2), 374–375, 375*f*, 413  
   ancestor's tale of mutations predisposing to inflammatory bowel disease, 376–377  
   identification of, 374  
   and innate immunity, 375–377  
     epidemiology of NOD2 in Crohn's disease, 376  
     knockout murine models, 375  
     mutations, 376–377  
     mutations and phenotype, 376  
   Nod2-deficient mice, 375  
 Nucleotide-binding domain (NBD), 375  
 Nutraceuticals and functional foods, 259  
 Nutrient withdrawal, 32  
 Nutrigenomics, 50
- O**  
 Obesity, 73, 79, 276–281  
   BMI, 74*t*  
   clinical cases, 74–77  
     family history, 74, 75*f*  
     follow-up, 77  
     genetic diagnosis, 75  
     patient history, 74  
     physical examination, 74  
   consequences, 73  
   family—obesity stigma, 86  
   genetic, 74  
     factors, 73  
   genetic counselling, 86  
   genetic/genomic factors, 279–281  
     environment and epigenetics/epigenomics, 280–281  
     family history and heritability, 279  
     metagenomics, 281  
     polygenic/multifactorial obesity, 280  
     rare monogenic diseases and syndromes, 279  
   molecular biology and pathophysiology, 79–83  
     energy balance, 79–80  
     energy expenditure, 79  
     energy intake, 79  
     fat storage, 79  
     leptin resistance, 81  
     leptin—melanocortin pathway, 81  
     monogenic obesity syndromes with intellectual disability, 81–83  
     nongenetic causes of obesity, 79  
   molecular diagnosis and therapy, 83–86  
     medication, 85–86  
   molecular systems underpinning clinical scenario, 77–78  
   monogenic nonsyndromic, 85  
 Obesity stigma, 86  
 Occludin, 20  
 Odds ratio (OR), 314–315, 427  
 17-OH. *See* 17-Hydroxyprogesterone (17-OH)  
 17-OH progesterone (17-OHP), 233  
 Oligogenic  
   inheritance, 293  
   traits, 533  
 Omega-3 fatty acids, 260  
 Omics technologies, 12–13, 13*f*, 438  
 OMIM database. *See* Online Mendelian Inheritance in Man database (OMIM database)  
 Omnigenic model, 373  
 Oncogene(s), 424  
   addiction, 497–499  
 “One-size-fits-all” model approach, 535  
 Online Mendelian Inheritance in Man database (OMIM database), 60, 66  
 ONT. *See* Oxford Nanopore Technology (ONT)  
 OR. *See* Odds ratio (OR)
- Oral antidiabetic drugs, 285–286  
 Oral medication of dystonia, 471  
 Organ transplantation, HLA and, 317–319  
   allorecognition, 317–318  
   AMR, 318, 318*f*  
   antibody-dependent cell-mediated cytotoxicity, 319  
   cellular rejection, 319  
   classification of rejection, 318  
   complement activation, 318–319  
   direct activation of endothelium, 319  
 ORMDL3 gene, 378  
 Orphan drugs, 505  
 OS. *See* Overall survival (OS)  
 Osimertinib, 498–499  
 Osteoporosis, 397–398  
 Ovalocytosis, 417  
 Ovarian cancer, 424–425  
 Ovarian development, 113  
 Overall survival (OS), 497–498  
 Overlapping clinical phenotypes, 439  
 Overweight, 276–281  
 Ovotesticular DSD, 116  
 Oxford Nanopore Technology (ONT), 59  
 Oxidation, 524  
 Oxidative phosphorylation, 43  
 Oxidative stress, 42–44, 454  
 Oxidative stress responses, 48  
 18-Oxocortisol, 181–182  
 OZI (ribozyme), 514
- P**  
 16p11.2 deletion, 78, 78*f*  
 p21 gene, 41  
 p450 oxidoreductase (POR), 111  
 p53. *See* Protein 53 (p53)  
 PA. *See* Primary hyperaldosteronism (PA); Propionic acidemia (PA)  
 PABP. *See* Polyadenylate-binding protein (PABP)  
 Pacific Biosciences (PACBIO), 59  
 PAI-1. *See* Plasminogen activator inhibitor-1 (PAI-1)  
 Paired-like homeobox 2B (PHOX2B), 371–372  
 Paired-like homeodomain 2 gene (*PITX2*), 195, 199  
 PAIS. *See* Partial androgen insensitivity syndrome (PAIS)  
 PAM. *See* Protospacer adjacent motif (PAM)  
 Pancreatic neuroendocrine tumors (pNET), 237  
 Pancreatic  $\beta$ -cell dysfunction, 284  
 Paneth cells, 375  
 Pantothenate kinase-associated neurodegeneration (PKAN), 481  
 Paracrine signaling, 18–19  
 Paradoxes with diseases, 380  
 Parafibromin, 225  
 Paraganglioma (PGL), 236–240  
   characteristics of gene mutations, 238*t*  
   clinical case, 236–240  
   molecular systems, 237–239  
   management, 240  
   protocol for lifetime surveillance, 240*t*  
   syndromes, 239*t*

- Parallels with diseases, 380  
 Parathyroid hormone (PTH), 222, 227  
 Parental genetic imprinting, 280–281  
 Parkin/PARK2, 463  
 Parkinson's disease (PD), 447, 451–452, 461–464  
   autosomal recessive inheritance, 463–464  
   epidemiology and pathophysiology, 461  
   genetics, 461–463  
   treatment, 464  
 Parkinsonism, 464  
 Paroxysmal kinesigenic dyskinesia.  
   *See* DYT10  
 PARP inhibitors. *See* Poly(ADP-ribose)  
   polymerase inhibitors (PARP inhibitors)  
 Partial androgen insensitivity syndrome  
   (PAIS), 112–113, 116  
 PAS. *See* PEG or alanine and serine (PAS)  
 Patent ductus arteriosus (PDA), 123  
 Pathogenesis of CKD, 439–440  
 Pathophysiological processes, 189  
 Pathophysiology of dystonia, 465  
 PCR. *See* Polymerase chain reaction (PCR)  
 PCSK9 inhibitors, 259  
 PCSK9. *See* Proprotein convertase subtilisin/  
   kexin type 9 (PCSK9)  
 PCT. *See* Proximal convoluted tubule (PCT)  
 PD. *See* Parkinson's disease (PD)  
 PD dementia (PDD), 452  
 PD-1. *See* Programed cell death-1 (PD-1)  
 PDA. *See* Patent ductus arteriosus (PDA)  
 PDD. *See* PD dementia (PDD)  
 PDE4D. *See* Phosphodiesterase 4D (PDE4D)  
 PEG. *See* Polyethylene glycol (PEG)  
 PEG or alanine and serine (PAS), 506  
 Pembrolizumab, 499–500  
 Peptide chain synthesis, 4–6, 6f  
 Peptidyl-tRNA, 34  
 Pericentriolar matrix, 39  
 Periventricular nodular heterotopia, 295  
 Permanent neonatal diabetes mellitus  
   (PNDM), 283  
 Peroxisomes, 44  
 Personalized medicine (PM), 523, 537–539,  
   538f  
   future of pharmacogenetics,  
     pharmacogenomics, and personalized  
     medicine, 533  
   with pharmacogenetics and  
     pharmacogenomics  
     clinical implementation of information,  
       528–530  
     clinical pharmacogenetics  
       implementation consortium, 530  
     ethical issues in genomic medicine, 532  
     genetic and molecular basis of  
       individual drug-response variation,  
       523–525  
   HGP, 522–523  
   pharmacogenomics and drug  
     development, 531–532  
   resources to collect and curate  
     pharmacogenetic variants, 532–533  
   tailoring or individualizing drug  
     therapy, 526–528  
     validation of genetic/genomic  
       information for drug response, 528  
 Personalized pharmacotherapy in CKD, 441  
 Personalized treatment, 85–86  
 PET. *See* Positron emission tomography  
   (PET)  
 “Petit mal”, 290  
 Peutz–Jeghers syndrome (PJS), 424  
 PFS. *See* Progression-free survival (PFS)  
 PGD. *See* Preimplantation genetic diagnosis  
   (PGD)  
 PGL. *See* Paraganglioma (PGL)  
 Ph chromosome. *See* Philadelphia  
   chromosome (Ph chromosome)  
 PHA2. *See* Pseudohypoadosteronism type 2  
   (PHA2)  
 Pharmacologic interaction mechanism (pi  
   mechanism), 313  
 Pharmaceutical and Medical Devices Agency  
   (PMDA), 532  
 Pharmacodynamics, 524  
   genetic factors in, 524–525  
 Pharmacogenetics, 521–522  
   clinical implementation of information,  
     528–530  
   cost of testing and turnaround time, 528  
   lack of strong evidence/weak evidence  
     base, 530  
   physician barriers, 529  
   SNP testing limitations in isolation,  
     528–529  
   specific protocols to guide decision-  
     making post–pharmacogenetic  
     testing, 530  
   stakeholder engagement including  
     regulator, 530  
   of epilepsy, 299–300  
     adverse antiepileptic drug reactions,  
       299–300  
     antiepileptic medication, 300  
     cytochrome enzymes, 300  
     drug response, 300  
     HLA, 299–300  
     hyponatremia, 300  
   future, 533  
   historical overview of, 522f  
   resources to collect and curate  
     pharmacogenetic variants, 532–533  
   sodium channel  
     blocking antiepileptic drugs, 300  
     genes, 300  
 Pharmacogenomics, 441, 521–522  
   clinical implementation of information,  
     528–530  
   cost of testing and turnaround time, 528  
   lack of strong evidence/weak evidence  
     base, 530  
   physician barriers, 529  
   SNP testing limitations in isolation,  
     528–529  
   specific protocols to guide decision-  
     making post–pharmacogenetic  
     testing, 530  
   stakeholder engagement including  
     regulator, 530  
   in clinical practice, 526–528  
   and drug development, 531–532  
   future, 533  
   historical overview, 522f  
 Pharmacokinetics, 523–524  
   genetic factors in, 524, 525f  
 Pharmacological agents, 99, 102–103  
   dystrophic epidermolysis bullosa, 103  
   Farber/spinal muscular atrophy with  
     progressive myoclonic epilepsy,  
       102–103  
   Joubert syndrome, 102  
   laminopathies, 103  
   thanatophoric dysplasia and  
     achondroplasia, 103  
 Pharmacological therapies, 164–165  
 Pharmacology, 50–51  
 Pharmacotherapy, 521  
 PharmGKB, 532–533  
 Phase 1 translation research (T1 research), 14  
 Phase 2 reentry phenomenon, 146–147  
 Phase 2 translation research  
   (T2 research), 14  
 Phase 3 translation research (T3 research), 14  
 Phase 4 translation research (T4 research), 14  
 Phenotype-driven approaches, 12  
 Phenotypic severity, 303  
 Phenotypically striking disorders, 301  
 Phenytoin, 300  
 Pheochromocytomas (PHEOs), 236–240  
   clinical case, 236–240  
   molecular systems, 237–239  
   gene mutation characteristics, 238f  
   management, 240  
   protocol for lifetime surveillance, 240f  
   syndromes, 239f  
 PHEOs. *See* Pheochromocytomas (PHEOs)  
 PHEX gene, 229  
 Philadelphia chromosome (Ph chromosome),  
   428–429, 497  
 Phosphatidylcholine, 17–18  
 Phosphatidylethanolamine, 17–18  
 Phosphatidylinositol, 17–18  
 Phosphatidylinositol 4,5-bisphosphate (PIP2),  
   19  
 Phosphatidylserine, 17–18  
 Phosphodiesterase 4D (PDE4D), 196  
   cAMP-specific gene, 196–197, 197f  
 Phosphoglycerides, 17  
 Phospholipase A2-associated  
   neurodegeneration (PLAN), 481  
 Phospholipid, 17  
 Phosphoric acid, 17  
 Phosphorylation, 19  
 PHOX2B. *See* Paired-like homeobox 2B  
   (PHOX2B)  
 PHP. *See* Pseudohypoparathyroidism (PHP)  
 Physician-diagnosed asthma, 354  
 Physiotherapy, 471  
 pi mechanism. *See* Pharmacologic interaction  
   mechanism (pi mechanism)  
 PIDs. *See* Primary immunodeficiency  
   diseases (PIDs)  
 Pilomatricomas, 425  
 PINK1. *See* PTEN-induced kinase 1 (PINK1)



- PIP2. *See* Phosphatidylinositol 4,5-bisphosphate (PIP2)
- piRNAs. *See* Piwi-associated RNAs (piRNAs)
- Pituitary adenoma, 217–221
  - genetics, 221
  - genetics and clinical characteristics of syndromes associating with, 222*t*
- PITX2*. *See* Paired-like homeodomain 2 gene (*PITX2*)
- Piwi-associated RNAs (piRNAs), 12
- PJS. *See* Peutz–Jeghers syndrome (PJS)
- PKA. *See* Protein kinase A (PKA)
- PKAN. *See* Pantothenate kinase-associated neurodegeneration (PKAN)
- PLA2G6* gene, 481
- Plain chest radiography, 157, 158*f*
- PLAN. *See* Phospholipase A2-associated neurodegeneration (PLAN)
- Plasma
  - membrane, 17–21
    - cell junctions, 19–21
    - cell signaling, 18–19
  - natriuretic peptides, 155
- Plasma renin activity (PRA), 171
- Plasminogen, 345
- Plasminogen activator inhibitor-1 (PAI-1), 345
- Platelets, inherited defects of, 349
  - Bernard–Soulier syndrome, 349
  - Glanzmann’s thrombasthenia, 349
  - inherited macrothrombocytopenia, 349
  - May–Hegglin anomaly, 349
  - storage pool disease, 349
  - Wiscott–Aldrich syndrome, 349
- Pluripotent stem cells (PSCs), 509
- PM. *See* Personalized medicine (PM)
- PMDA. *See* Pharmaceutical and Medical Devices Agency (PMDA)
- PNDM. *See* Permanent neonatal diabetes mellitus (PNDM)
- pNET. *See* Pancreatic neuroendocrine tumors (pNET)
- POF. *See* Premature ovarian failure (POF)
- Polar microtubules, 41
- POLII. *See* Polymerase II (POLII)
- Poly-A polymerase, 30
- Poly(ADP-ribose) polymerase inhibitors (PARP inhibitors), 424–425, 499
- Polyadenylate-binding protein (PABP), 32
- Polycystin 2, 42
- Polyethylene glycol (PEG), 505–506
- Polygenic granulomatous diseases, 380
- Polygenic/multifactorial obesity, 280
- Polymerase chain reaction (PCR), 342–343
- Polymerase II (POLII), 24–28
- Polymorphisms, 9, 189, 407–408, 524
  - in 17q21 chromosome, 357
  - immunity-related gene polymorphisms, 411
    - NRAMP1 gene polymorphisms, 413
    - single-nucleotide polymorphisms, 127–128
- Polypeptide chains, 349
- Polyuria, 275
- POMC. *See* Proopiomelanocortin (POMC)
- POMC gene, 85–86
- Populations, 9
  - data, 66
  - genetic studies, 354
- POR. *See* p450 oxidoreductase (POR)
- Positional candidate-gene approach, 374
- Positional cloning, 357, 359
- Positron emission tomography (PET), 157, 450
- “Poster child” of pharmacogenomics, 526
- Post-genomic analysis, 47
- Postmortem tissue analysis, 470
- Posttranslational influence, 316
- PRA. *See* Plasma renin activity (PRA)
- Prader–Willi syndrome (PWS), 81–83, 85
- pRB (retinoblastoma tumor suppressor protein family), 41
- Precision genome editing, 518
- Precision medicine, 523, 535–539, 538*f*
- Precursor globin mRNA molecule, 329
- Prediabetes, 275–276
- Predictive, personalized, preemptive, and participatory of medicine (4 Ps of medicine), 537–538
- Predictive genetic testing for cancer, 431–434
- Predictive testing of family members, 163
- Pregnancy, 321
  - amiloride during, 178
  - CH during, 217
  - HLA sensitization, 321
  - VTE during, 344
- Preimplantation genetic diagnosis (PGD), 331–332
- Premature conduction disorders, 156–157
- Premature ovarian failure (POF), 110–111
- Premature termination codons (PTCs), 102
- Prenatal diagnosis, 331
  - noninvasive, 332
- Preprescription genetic information, 526–527
- Presenilin-1 (PSEN1), 63
- Pretangles, 452–453
- Previous organ transplantation, 321
- Primaquine, 524
- Primary cilium, 41–42
- Primary gonadal failure, 240–242
- Primary hyperaldosteronism (PA), 230–232
  - clinical case, 230–231
  - molecular systems, 231
  - management of patients with, 232
- Primary hyperparathyroidism, 222–226
  - clinical case, 222
  - management of patients with, 225–226
    - medical management, 225–226
    - surgical management, 225
  - molecular systems, 223–225
- Primary hypertension, 169
- Primary immunodeficiency diseases (PIDs), 99
- Primary lipoprotein disorders, treatment of, 256
  - LDL-C, 258*t*
  - rare genetic lipid disorders, 257*t*
- Primitive cardiac tube, 121–122
- Primordial germ cells, 107
- Prion diseases, 448–449
- Proband sequencing, 61
- Programed cell death-1 (PD-1), 499
- Progression-free survival (PFS), 497–498
- Progressive myoclonic epilepsies, 299, 302
- Proinflammatory genes, 204
- Prokaryotes, 4
- Prokineticin 2 (PROK2), 221
- Prolactin, 217
- Proline-rich transmembrane protein 2 (PRRT2), 470
- Prometaphase, 41
- “Proof of principle” studies, 370, 514–515
- Proopiomelanocortin (POMC), 80–81
- Propionic acidemia (PA), 48
- Proprotein convertase subtilisin/kexin type 9 (PCSK9), 38
- Proteasome, 39
- Protein 53 (p53), 41, 408
- Protein C, 344
  - deficiency, 344
- Protein kinase A (PKA), 19
- Protein kinase C, 19
- Protein(s), 6, 18
  - biomarkers, 438
  - deglycase, 463
  - drugs, 505–506
  - modulation, 99, 102
  - protein S deficiency, 344
  - synthesis, 32–34
    - ribosomes and mRNA translation, 32–34
  - turnover and cell size control, 38–39
  - tyrosine phosphatase, 125
- Proteomics, 12–13, 48, 374
- Prothrombin allele G20210A, 344
- Proto-oncogenes, 526
- Protons (<sup>1</sup>H), 51–52
- Protospacer adjacent motif (PAM), 517
- Proximal convoluted tubule (PCT), 169–170
- PRRT2*. *See* Proline-rich transmembrane protein 2 (*PRRT2*)
- PSCs. *See* Pluripotent stem cells (PSCs)
- PSEN1. *See* Presenilin-1 (PSEN1)
- Pseudohypoadosteronism type 2 (PHA2), 171
- Pseudohypoparathyroidism (PHP), 227
  - management of patients with, 228
- PTCs. *See* Premature termination codons (PTCs)
- PTEN* gene, 424
- PTEN-induced kinase 1 (PINK1), 463
- PTH. *See* Parathyroid hormone (PTH)
- PTPN11 mutations, 125
- PTPN2 gene. *See* T-cell protein tyrosine phosphatase gene (PTPN2 gene)
- PTPN22* gene, 272
- Public health promoters, 393
- PubMed, 67
- Pulmozyme, 508
- Pulsatile GnRH therapy, 221
- Purified omega-3 fatty acids, 260
- Purkinje system branches, 145
- PWS. *See* Prader–Willi syndrome (PWS)
- Pyrosequencing, 58
- Pyruvate kinase type M2 (M2-PK), 49–50



- Q**  
 22q11.2 deletion, 487–488  
 QT interval, 138  
 QT<sub>c</sub>. *See* Corrected QT interval (QT<sub>c</sub>)  
 Qualitative defects of Hbs, 334–336  
   defects of erythroid heme biosynthesis, 336  
   HbC, 335  
   HbE, 335  
   HbM, 335–336  
   high-oxygen affinity Hb, 336  
   low-oxygen affinity Hb, 336  
   sickle-cell Hb, 334–335  
   unstable Hb, 336  
 Quiescent memory B cells, 318  
 Quintessential autoimmune disease, 380
- R**  
 “R-on-T” phenomenon, 143  
 RA. *See* Rheumatoid arthritis (RA)  
 RAAS system. *See* Renin–angiotensin aldosterone system (RAAS system)  
 Rabson–Mendenhall syndrome, 284  
 Raman spectroscopy, 51  
 Ran-binding protein 2 gene (RANBP2 gene), 412  
 Randomized controlled study designs, traditional, 532  
 Randomized controlled trials (RCTs), 256  
 Ras of complex (Roc), 462–463  
 Rasburicase, 524  
 RAS–mitogen-activated protein kinase (RAS–MAPK), 94  
   syndromes, 425  
 RASopathies, 94, 95*t*, 125  
 Raynaud’s disease, 393–394  
 RBBP8. *See* Retinoblastoma binding protein 8 gene (RBBP8)  
 RBM12. *See* RNA-binding-motif protein 12 (RBM12)  
 RCTs. *See* Randomized controlled trials (RCTs)  
 RDEB. *See* Recessive dystrophic epidermolysis bullosa (RDEB)  
 Reactive oxygen species (ROS), 43, 454  
   ROS1 translocations, 498  
 Recessive dystrophic epidermolysis bullosa (RDEB), 103  
 Recessive inheritance, 354  
 Recombinant DNA technology, 7, 505–506  
 Recombinant pharmacotherapy, 505–506, 506*t*  
 Recombinant protein drugs and vaccines, 505–508  
   recombinant pharmacotherapy, 505–506  
   recombinant therapeutic enzymes, 508, 508*t*  
   recombinant vaccines, 506–508  
     hepatitis B recombinant vaccine, 507  
     HIV vaccines, 507–508  
     HPV vaccine, 507  
 Recombivax, 507  
 Recurrent chromosomal fusion events, 500  
 Recurring theme, 293  
 Reductive approaches, 367  
 Refractory hypertension, 232–233  
 Relative risk (RR), 353, 423  
 Renal outer medullary potassium (K) channel (ROMK), 173–174  
 Renal tubular defects, 235–236  
 Renin–angiotensin aldosterone system (RAAS system), 164, 169–170  
 Repeat expansion disorders, 302  
 Replication cohorts, 361  
 Repolarization theories, 146–147  
 RER. *See* Rough ER (RER)  
 Resting electrocardiography, 156  
 Retinoblastoma binding protein 8 gene (RBBP8), 207  
 Reverse transcriptase PCR (RT-PCR), 342–343  
 RF. *See* Rheumatoid factor (RF)  
 Rheumatoid arthritis (RA), 312, 380, 395–400  
   articular features, 396–398  
   clinical subtypes, 395–396  
   genetic factors, 395  
   histological appearance of synovium in, 397*f*  
   molecular pathology, 395  
   nonarticular manifestations, 397*t*, 398–399  
   nonarticular systemic manifestations in, 398*f*  
   pathogenesis, 396*f*  
   SLE, 400  
   treatment, 400  
 Rheumatoid factor (RF), 395  
   transient production of, 395  
 Ribonucleic acid (RNA), 4, 475  
   genes in human genome, 13*t*  
   RNA POLII, 30  
   RNA-Seq, 373–374  
 Ribosomal RNAs (rRNAs), 11, 30  
 Ribosome, 30, 32–34  
   ribosome–translocon association, 34–35  
 Ribozymes, 514  
 RISC. *See* RNA-induced silencing complex (RISC)  
 RNA. *See* Ribonucleic acid (RNA)  
 RNA interference (RNAi), 99, 505, 514–515  
   and cancer, 426  
   gene-silencing technology, 515  
 RNA-binding-motif protein 12 (RBM12), 492  
 RNA-induced silencing complex (RISC), 514  
 RNAi. *See* RNA interference (RNAi)  
 Roc. *See* Ras of complex (Roc)  
 Rolandic epilepsy, 290  
 Romano–Ward syndrome, 138–139  
 ROMK. *See* Renal outer medullary potassium (K) channel (ROMK)  
 ROS. *See* Reactive oxygen species (ROS)  
 Rough ER (RER), 34  
 Roux-en-Y gastric bypass, 85  
 RPE65 gene, 100  
 RR. *See* Relative risk (RR)  
 RREB1 gene, 129  
 rRNAs. *See* Ribosomal RNAs (rRNAs)  
 RT-PCR. *See* Reverse transcriptase PCR (RT-PCR)  
 RTS,S (malaria vaccine), 419  
 Rubinstein–Taybi syndrome, 98*t*  
 Runt-related transcription factor 1 (RUNX-1), 427–428  
 Ryanodine receptor (RyR), 136–137  
   RyR2, 145  
 RyR. *See* Ryanodine receptor (RyR)  
 RYR2 mutation, 134, 147
- S**  
 602S allele, 413  
 SAB assay. *See* Single-antigen bead assay (SAB assay)  
 SAIDs. *See* Systemic autoinflammatory disorders (SAIDs)  
*Salmonella typhimurium*, 375  
 Salt-sensitive hypertension, 441  
 Sanger sequencing, 58, 346, 348  
 Sarcoidosis, 393–394  
 Sarcolemma, 136  
 Sarcomeres, 136  
 Sarcoplasmic reticulum (SR), 135–137  
 SASP. *See* Senescence-associated secretory phenotype (SASP)  
 Say–Barber–Biesecker syndrome/  
   Young–Simpson syndrome, 98*t*  
 SBP. *See* Systolic BP (SBP)  
 SC35 protein, 125  
 SCA. *See* Spinocerebellar ataxias (SCA)  
 Schizophrenia (SZ), 487, 492  
 Schwartz score, 138, 139*t*  
 SCID. *See* Severe combined immunodeficiency (SCID)  
 SCLC. *See* Small-cell lung cancer (SCLC)  
 SCN4B protein, 141  
 SCN5A gene, 147, 162–163  
 SCNN1A gene, 176  
 SCNN1B gene, 176  
 SCNN1G gene, 176  
 SCORE. *See* Systematic coronary risk estimation (SCORE)  
 SDH. *See* Succinate dehydrogenase (SDH)  
 SEC24C gene, 129  
 Second-generation sequencing, 58–59  
 Secondary hemostasis congenital disorders, 341  
 Secondary hyperparathyroidism, 228  
 Secondary skin cancers, 499  
 Secretory pathways, 34–38, 35*f*  
 Segawa’s disease, 469  
 Segregation analysis, 354–355, 370  
 Seizures, 289–290  
   complex partial, 290  
   diagnosis of, 289  
   focal to bilateral tonic-clonic, 290  
   generalized tonic-clonic, 290  
   partial onset, 290  
   predisposition, 290  
 Selectins, 21  
 SELEX. *See* Systemic Evolution of Ligands by EXponential enrichment (SELEX)  
 Self-surveillance complex, 309–310  
 SEMA3E genes, 89  
 Semidirect pathway, 317–318  
 Semiology, 289–290  
 Semiquantitative, hierarchical evidence-based rules for locus interpretation system (SHERLOC system), 68

- Senescence, 393  
 Senescence-associated secretory phenotype (SASP), 393  
 Senescent cells, 393  
 Sensitization, 321  
 Sepsis, 390–391  
 Sequence alignment, 65  
 Serine- and arginine-rich proteins (SR proteins), 29  
   heterodimerization, 34–35  
 SERKAL syndrome, 113  
 Seronegative RA, 395  
 Serum creatinine, 437–438  
 Severe combined immunodeficiency (SCID), 99  
 Severe hyperphagia, 75  
 Severe myoclonic epilepsy of infancy, 293  
 Sex chromosome disorder of sex  
   development, 108–110  
 Sex chromosomes, 4  
 Sex determination, 107  
 Sex determining region Y-box 10 (SOX10), 221  
 SGCE gene. *See* Epsilon-sarcoglycan gene (SGCE gene)  
 SGCE mutations. *See* DYT11  
 SGLT2. *See* Sodium–glucose cotransporter (SGLT2)  
 Shared epitope, 312  
 SHFM. *See* Split hand–foot malformation (SHFM)  
 SHIVA trial, 500–501  
 Short QT syndrome (SQTS), 137, 144, 148.  
   *See also* Long QT syndrome (LQTS)  
   clinical risk assessment and therapy in, 148  
   generation and propagation of arrhythmia, 144  
   molecular basis and features, 144f  
 Short-acting beta 2 agonists, 363  
 “Short–long–short” sequence, 143  
 SHP2 mutants, 125  
 Sickle cell  
   anemia, 335  
   disease, 10, 101  
   Hb, 334–335  
 Sickled erythrocytes, 334–335  
 Sickling syndrome diagnosis, 335  
 Signal recognition particle (SRP), 34  
 Signal-to-noise ratio, 50–51  
 Single dysregulated protein, 497  
 Single gene test, 301  
 Single-antigen bead assay (SAB assay), 319–320  
 Single-gene disorders, 124–126, 125f  
   associated with stroke, 189–191, 190f  
   isolated congenital heart disease causing  
     by single gene, 126  
   RASopathies, 125  
   transcription factor–related disorders, 125–126  
 Single-gene mutation, 252, 403–404  
 Single-nucleotide polymorphism (SNPs), 9, 67, 77, 122, 127–128, 171, 192, 341, 359, 370–371, 394, 403–404, 427, 438, 523  
   microarrays, 487  
   SNP rs1801274, 409  
   SNP-array analysis, 78  
   testing limitations in isolation, 528–529  
 Single-nucleotide variants (SNVs), 492  
 Single-stranded DNA molecules, 59  
 Single-stranded oligodeoxynucleotide (ssODN), 518  
 SIR. *See* Standardized incidence ratios (SIR)  
 siRNA. *See* Small interfering RNA (siRNA)  
 SIRS. *See* Systemic inflammatory response syndrome (SIRS)  
 SJS. *See* Stevens–Johnson syndrome (SJS)  
 SLE. *See* Systemic lupus erythematosus (SLE)  
 Small gene panels, 60–61  
 Small interfering RNA (siRNA), 514  
 Small nuclear ribonucleoprotein particles (snRNP), 29  
 Small nuclear RNA (snRNA), 29  
 Small nucleolar RNAs (snoRNA), 29  
 Small oligomeric aggregates, 454  
 Small-cell lung cancer (SCLC), 523  
 Small-vessel disease, 190  
 SMA–PME. *See* Spinal muscular atrophy–progressive myoclonic epilepsy (SMA–PME)  
 Smith–Fineman–Myers syndrome, 90–91  
 SMN. *See* Survival of motor neuron 1 (SMN)  
 SMN1 gene, 29  
 SNAP25. *See* Synaptosomal-associated protein 25 (SNAP25)  
 SNARE. *See* Soluble attachment receptor (SNARE)  
 SNCA/PARK1, 461–462  
 SNORD116 deficiency, 81–82  
 snoRNA. *See* Small nucleolar RNAs (snoRNA)  
 SNPs. *See* Single-nucleotide polymorphism (SNPs)  
 snRNA. *See* Small nuclear RNA (snRNA)  
 snRNP. *See* Small nuclear ribonucleoprotein particles (snRNP)  
 SNTA1 protein, 141  
 SNVs. *See* Single-nucleotide variants (SNVs)  
 SOD. *See* Superoxide dismutase (SOD)  
 SOD1 gene, 43  
 Sodium channel  
   blocking antiepileptic drugs, 300  
   genes, 300  
 Sodium–glucose cotransporter (SGLT2), 275–276  
 Soluble attachment receptor (SNARE), 36  
 Somatic gene therapy, 511–512  
 South Asian ovalocytosis, 417  
 SOX10. *See* Sex determining region Y-box 10 (SOX10)  
 SOX17. *See* SRY-box 17 gene (SOX17)  
 SOX3 genes, 111  
 SOX8 genes, 111  
 Sphingolipids, 17–18  
 Sphingomyelin, 17–18  
 Spiegelmers, 516  
 Spinal muscular atrophy–progressive myoclonic epilepsy (SMA–PME), 102–103  
 Spinocerebellar ataxias (SCA), 472  
   clinical and genetic description, 472f  
   with epilepsy, 299  
 SCA2, 476–478  
 SCA3, 478  
 SCA7, 478–479  
 Spirometry, 353  
 Spliceopathies, 97  
 Splicing, 30  
 Split hand–foot malformation (SHFM), 93–94, 94f  
 SQTS. *See* Short QT syndrome (SQTS)  
 SR. *See* Sarcoplasmic reticulum (SR)  
 SR proteins. *See* Serine- and arginine-rich proteins (SR proteins)  
 SRD5A2 gene, 112, 116–117  
 SRNS. *See* Steroid-resistant nephrotic syndrome (SRNS)  
 SRP. *See* Signal recognition particle (SRP)  
 SRY gene, 107, 110, 116  
 SRY-box 17 gene (SOX17), 207  
 ssODN. *See* Single-stranded oligodeoxynucleotide (ssODN)  
 Stabilized microtubules, 41  
 Stakeholder engagement, 530  
 Standardized incidence ratios (SIR), 292, 292f  
*Staphylococcus*, 362  
 StAR protein. *See* Steroidogenic acute regulatory protein (StAR protein)  
 StAR-related lipid transfer domain containing 13 gene (STARD13), 207  
 STAT3 function, 415–416  
 Statins, 256  
 Stem cell  
   stem cell–based therapy, 447  
   therapy, 509–511  
   transplant, 335  
 Sterile inflammation, 392  
 Steroid hormones, 18–19  
 Steroid synthesis defects, 114–115, 115f  
 Steroid-resistant nephrotic syndrome (SRNS), 439  
 Steroidogenic acute regulatory protein (StAR protein), 111  
 Stevens–Johnson syndrome (SJS), 525, 527  
 sTFs. *See* Tissue-specific TFs (sTFs)  
 Storage Pool Disease, 349  
 Stratified medicine, 536, 539–541  
   criteria for, 540f  
*Streptococcus*, 362  
 Stroke, molecular basis of, 189  
   clinical implication, 207  
   genetics of, 189  
   forms of, 191–207  
   single-gene disorders associated with, 189–191  
 Structural X chromosome, 108–109  
 Subarachnoid hemorrhage, 189  
   delayed cerebral ischemia after, 204f  
   molecular pathophysiology of, 202–204  
   etiology, 202f  
   intracranial aneurysm, 203f  
   molecular genetics of, 205–207  
 Subcutaneous nodules, 397–398  
 Substantia nigra, 451–452  
 Succinate dehydrogenase (SDH), 236, 238

Sudden cardiac death, 158–159  
 Sulphonamides, 524  
 Superoxide ( $O_2^-$ ), 43  
 Superoxide dismutase (SOD), 43  
 Surface ECG, 138  
   relation of action potential to, 135  
 Surgery, 259–260  
 Survival of motor neuron 1 (SMN), 29  
 “Survivor guilt” phenomenon, 163–164  
 Svedberg coefficients (S coefficients), 30  
 Symmetrical small-joint arthropathy, 377–378  
 Sympathetic nervous system, 148, 170  
 Symptomatic control, 363  
 Synaptic neurotransmission, 475  
 Synaptosomal-associated protein 25 (SNAP25), 470  
 Syndrome, 89  
   with genetic connections, 94–96  
 Syndromic obesity, 74, 81–82  
 Synthetic lethality, 499  
 Systematic coronary risk estimation (SCORE), 249  
 Systemic autoinflammatory disorders (SAIDs), 394  
 Systemic Evolution of Ligands by EXponential enrichment (SELEX), 516  
 Systemic hypersensitivity reactions, 299–300  
 Systemic hypertension, 169  
   molecular pathophysiology  
     BP regulation, 169–170  
     clinical implications from genetic studies, 184  
     familial hyperaldosteronism, 181–183  
     future perspectives, 184  
     genetic overlap of monogenic and essential hypertension, 183–184  
     genetics of hypertension, 170–171  
     immune system, 170  
     monogenic forms of, 171–181  
     natriuretic peptides, 170  
     RAAS system, 170  
     sympathetic nervous system, 170  
 Systemic inflammatory response syndrome (SIRS), 390  
 Systemic lupus erythematosus (SLE), 314–315, 380, 393–394, 399f, 400  
 Systemic neutralization of cytokines, 393  
 Systems biology, 12–13  
 Systolic BP (SBP), 169  
 SZ. *See* Schizophrenia (SZ)

## T

T wave, 138  
   alternans, 143–144  
   inversion, 156  
 T-cell protein tyrosine phosphatase gene (PTPN2 gene), 371–372  
 T-cell regulation and function, 407  
 t-SNARE. *See* Targeted destination soluble attachment receptor (t-SNARE)  
 T1DM. *See* Type 1 diabetes mellitus (T1DM)  
 T2DM. *See* Type 2 diabetes mellitus (T2DM)  
 TA-GVHD. *See* Transfusion-associated GVHD (TA-GVHD)  
 Tachykinin family  
   TAC3 and TACR3, 221  
 Tailoring drug therapy, 526–528  
 TALENs. *See* Transcription activator-like effector nucleases (TALENs)  
 Tamm–Horsfall protein, 440  
 Tandem-MS analysis, 49  
 Targeted clinical management, 441  
 Targeted clinical surveillance, 441  
 Targeted destination soluble attachment receptor (t-SNARE), 36  
 Targeted gene panel analysis in CKD, 440–441  
 Targeted molecular diagnosis and therapy for cancer, 497  
   histology agnostic treatment, 499–500  
   making common cancer rare, 501  
   molecularly targeted therapy limitations  
     in cancer, 500–501  
   oncogene addiction, 497–499  
   synthetic lethality, 499  
   for cardiac arrhythmias, 147–149  
     Brugada syndrome, 149  
     CPVT, 148–149  
     long QT syndrome, 147–148  
     short QT syndrome, 148  
   of dystonia, 465  
   of obesity disorders, 83–86  
     copy number variation, 83  
     DNA diagnostics, 83  
     leptin, 83  
   Targeted oncoproteins, 498–499  
 TATA-binding protein (TBP), 24–28  
 Tau protein, 452–453, 458  
 TBMD. *See* Thin basement membrane disease (TBMD)  
 TBP. *See* TATA-binding protein (TBP)  
 TBX5 gene, 122  
 TBX5, 125–126  
 TCF7-L2. *See* Transcription factor-7-like 2 (TCF7-L2)  
 tCho. *See* Total choline-containing compounds (tCho)  
 TCOF1 gene, 30  
 TDFG1 genes, 92  
 TDI. *See* Thanatophoric dysplasia type I (TDI)  
 TDP-43. *See* TAR DNA-binding protein 43 (TDP-43)  
 TEN. *See* Toxic epidermal necrolysis (TEN)  
 Tendon xanthomata, 253–254  
 Tenosynovitis of flexor tendons, 397–398  
 Termination of protein synthesis, 34  
 Testicular (gonadal) development, disorders of, 110–111  
 Tethering factors, 36  
 Tetrahydrobiopterin (TH4), 469  
 Tetralogy of Fallot (TOF), 121  
 Tetramer, 22  
 Tetrameric molecule, 327  
 TF. *See* Transcription factor (TF)  
 TFIIB, 24–28  
 TFIIF, 24–28  
 TFIIH, 24–28  
 TFIIs, 24–28  
 TGFB3. *See* Transforming growth factor beta 3 (TGFB3)  
 TGFbeta, 425  
 TH4. *See* Tetrahydrobiopterin (TH4)  
 Thalassemia intermedia (TI), 328, 333–334  
   patients with, 328–329  
 Thalassemia major (TM), 328  
   patients with, 328–329  
 Thalassemias, 10, 328  
    $\alpha$  thalassemias, 332–334  
    $\beta$  thalassemias, 328–329  
   classification of, 328–330  
   genetics of, 328–330  
 Thanatophoric dysplasia and achondroplasia, 103  
 Thanatophoric dysplasia type I (TDI), 103  
 THAP1 gene, 465  
 THAP1 mutations. *See* DYT6  
 Thin basement membrane disease (TBMD), 439  
 Third-generation sequencing, 59  
 Three prime repair exonuclease 1 gene (TREX1), 190  
 3D facial image analysis, 90  
 Thrombosis, 341, 343–345  
   case study, 343  
   genetic basis of, 343–345  
   molecular basis of causes, 343–345  
     APCR, 344  
     ATIII deficiency, 343  
     FVL, 344–345  
     hyperhomocysteinemia, 345  
     inherited deficiency of fibrinolysis, 345  
     Protein C deficiency, 344  
     Protein S deficiency, 344  
     molecular genetics of APLS, 343  
 Thymic stromal lymphopoietin (TSLP), 360  
 Thymine (T), 4  
 Thyrotropin-stimulating hormone (TSH), 217  
 TI. *See* Thalassemia intermedia (TI)  
 Tight junctions, 20  
 Timothy syndrome, 141  
 Tissue-specific stem-cell transplantation, 509  
 Tissue-specific TFs (sTFs), 28  
 Tissue-type plasminogen activator (tPA), 345  
 TKIs. *See* Tyrosine kinase inhibitors (TKIs)  
 TLR. *See* Toll-like receptor (TLR)  
 TM. *See* Thalassemia major (TM)  
 TMD. *See* Transient myeloproliferative disorder (TMD)  
 tNAA. *See* Total N-acetyl aspartate (tNAA)  
 TNDM. *See* Transient neonatal diabetes mellitus (TNDM)  
 TNF- $\alpha$ . *See* Tumor necrosis factor alpha (TNF- $\alpha$ )  
 TNFRSF13B. *See* Tumor necrosis factor receptor superfamily, member 13B gene (TNFRSF13B)  
 TNFSF15 gene. *See* Tumor necrosis factor ligand superfamily member 15 gene (TNFSF15 gene)  
 TNNT2. *See* Troponin I (TNNT2)  
 TNNT2 gene. *See* Troponin T gene (TNNT2 gene)  
 TOF. *See* Tetralogy of Fallot (TOF)  
 Toll-like receptor (TLR), 403–404  
   TLR1, 412

- Toll-like receptor (TLR) (*Continued*)  
 TLR2, 412  
 TLR3, 514–515  
*TorsinA* mutations. *See* DYT1  
 Total choline-containing compounds (tCho), 49–50  
 Total *N*-acetyl aspartate (tNAA), 457–458  
 Toxic epidermal necrolysis (TEN), 394, 527  
*TP53* gene, 41, 408  
 tPA. *See* Tissue-type plasminogen activator (tPA)  
 tracrRNA. *See* Transactivating crRNA (tracrRNA)  
 TRALI. *See* Transfusion-related acute lung injury (TRALI)  
 Trametinib, 499  
*Trans* e-QTL, 311  
*Trans* Golgi membrane, 37  
 Transactivating crRNA (tracrRNA), 517  
 Transactive response (TAR) DNA-binding protein 43 (TDP-43), 453  
 Transcription activator-like effector nucleases (TALENs), 517  
 Transcription factor (TF), 19  
 Transcription factor-7-like 2 (TCF7-L2), 273–274  
 Transcription factor–related disorders, 125–126  
   CHARGE association, 125–126  
 Transcription repressors, 28  
 Transcriptome, 373–374  
 Transcriptome sequencing, 69  
 Transcriptomics, 12–13, 47, 373–374  
 Transcriptomopathies, 96–97, 97*t*  
 Transfer RNA (tRNAs), 11, 32–33  
   tRNA<sup>Met</sup>, 32–33  
 Transferrin receptor, 38  
 Transforming growth factor beta 3 (TGFB3), 161–162  
 Transfusion-associated GVHD (TA-GVHD), 322–323  
 Transfusion-related acute lung injury (TRALI), 322  
 Transient immune response, 390  
 Transient myeloproliferative disorder (TMD), 429–430  
 Transient neonatal diabetes mellitus (TNDM), 283  
 Translational genome research, 14  
 Translational human genomics, 14  
 Translocon, 34–35  
 Transmembrane channel abnormalities, 475  
 Transmission of hereditary factors, 3–4  
 Transplantation  
   kidney, 442  
   liver, 259–260  
   low *vs.* high expression mismatches in, 315  
   organ, 317–319  
   tissue-specific stem-cell, 509  
 Transposase-accessible chromatin sequencing (ATAC-seq), 378  
 Transthoracic echocardiography, 158  
 Trastuzumab, 526  
 Trauma, 390–391  
*TRDN* gene, 145  
 Treacher Collins syndrome, 30  
*TREX1*. *See* Three prime repair exonuclease 1 gene (*TREX1*)  
 Trigger factors, 202–203  
 Trihexyphenidyl, 471  
 Trinucleotide repeat disorders, 475  
 Trio sequencing, 61  
 Triplet repeat expansions, 60  
 TRK. *See* Tropomyosin receptor kinase (TRK)  
 tRNAs. *See* Transfer RNA (tRNAs)  
*Tropheryma whipplei*, 415  
 Tropomyosin, 136–137  
 Tropomyosin receptor kinase (TRK), 499  
 Troponin I (TNNI3), 161–162  
 Troponin T gene (TNNT2 gene), 63, 161–162  
 TruSight One Expanded featuring 6794 genes, 61  
 TruSight One featuring 4813 genes, 61  
 TSG. *See* Tumor suppressor gene (TSG)  
 TSH. *See* Thyrotropin-stimulating hormone (TSH)  
 TSLP. *See* Thymic stromal lymphopoietin (TSLP)  
 Tuberculin skin test (TST), 412  
 Tuberculosis, 412  
 Tumor  
   heterogeneity, 501  
   markers in circulating blood, 430  
 Tumor necrosis factor alpha (TNF- $\alpha$ ), 391, 409  
 Tumor necrosis factor ligand superfamily member 15 gene (TNFSF15 gene), 371  
 Tumor necrosis factor receptor superfamily, member 13B gene (TNFRSF13B), 205–207  
 Tumor suppressor gene (TSG), 424  
 Turner syndrome, 380  
 Twin genetic studies, 355–356, 356*t*, 369–370  
 Two-allele gene model, 354  
 Two-hit hypothesis, 424  
 Type 1 diabetes mellitus (T1DM), 270, 312, 371–372  
   autoimmunity and, 272–273  
   HLA and, 272  
   molecular mechanisms in, 270–272  
 Type 1A PHP, 227  
 Type 1B PHP, 227  
 Type 2 diabetes mellitus (T2DM), 10, 273–274, 380  
   epidemiology, 273  
   genetic factors in, 273–274  
 Type 2B VWD, 347  
 Type 2M VWD, 347  
 Type 3 hyperlipoproteinemia, 256  
 Tyrosine kinase inhibitors (TKIs), 497, 523
- U**  
 U waves, 138  
 Ubiquitin, 39  
 Ubiquitin-specific protease gene (USP3), 161  
 UC. *See* Ulcerative colitis (UC)  
 UCaCr. *See* Urinary calcium:creatinine (UCaCr)  
 UK 100, 000 Genomes Project, 539, 541, 542*f*  
 UK Genetic Testing Network (UKGTN), 60  
 Ulcerative colitis (UC), 367  
 Ultrarapid metabolizers, 524  
 Unc5-like kinase 1 (ULK1), 38  
 Unfolded protein response (UPR), 36  
 Uniparental isodisomy/somatic deletion of  $\beta$ -globin gene, 331  
 United Kingdom National Heart Failure Audit, 153  
 Unstable Hb, 336  
 Unverricht–Lundborg disease, 295, 302  
 UPR. *See* Unfolded protein response (UPR)  
 Upstream deletions, 330  
 Uracil (U), 4–6  
 Urinary calcium:creatinine (UCaCr), 224–225  
 Uromodulin, 440  
 US Food and Drug Administration (US FDA), 100, 505, 521, 532  
 USP3. *See* Ubiquitin-specific protease gene (USP3)
- V**  
 V-SNARE. *See* Vesicle soluble attachment receptor (V-SNARE)  
 VACTERL syndrome, 128  
 Valproate, 299  
 Variant hemoglobins, 327  
 Variants, clinical interpretation of, 66–69  
   ACMG guidelines, 68  
   allelic data, 68  
   computational and predictive data, 66–67  
   de novo status and segregation data, 67–68  
   functional data, 67  
   future developments in clinical genomics, 68–69  
   other data, 68  
   other databases, 68  
   population data, 66  
 Variants of uncertain significance (VUS), 60, 63  
 Vascular endothelial growth factor (VEGF), 127–128  
   VEGF allele +405C, 127–128  
 Vascular system, structural defects of, 348  
   EDS, 348  
   hereditary hemorrhagic telangiectasia, 348  
 Vasopressin, 458  
 VEGF. *See* Vascular endothelial growth factor (VEGF)  
 Velocardiofacial syndrome. *See* 22q11.2 deletion  
 Vemurafenib, 497–498, 500  
 Venous thromboembolism (VTE), 343  
 Ventricular septal defect (VSD), 126  
 Very low-density lipoprotein (VLDL), 245–246  
 Vesicle soluble attachment receptor (V-SNARE), 36  
 Vesicular trafficking, 34–38  
 VHL disease. *See* Von Hippel–Lindau disease (VHL disease)  
 Vimentin, 320–321  
 Vitamin D and DM, 281–282  
 Vitamin K epoxide reductase complex genes (*VKORC1* genes), 524, 526–527



VLDL. *See* Very low-density lipoprotein (VLDL)  
 Von Hippel–Lindau disease (VHL disease), 236, 238, 426  
 Von Willebrand disease (VWD), 346–348, 347f  
   type 1 VWD, 347f  
 Von Willebrand factor (VWF), 346  
 VPS35/PARK17, 463  
 VSD. *See* Ventricular septal defect (VSD)  
 VTE. *See* Venous thromboembolism (VTE)  
 VUS. *See* Variants of uncertain significance (VUS)  
 VWD. *See* Von Willebrand disease (VWD)  
 VWF. *See* Von Willebrand factor (VWF)

## W

Warfarin, 524  
   use as anticoagulant, 526–527  
 Watson–Crick base pairing, 512–513  
 Watson–Crick model, 4, 5f  
 WD repeat domain phosphoinositide-interacting protein 4 (WIP14), 482  
 Weight loss, 275  
 Weight maintenance, 73  
 WES. *See* Whole-exome sequencing (WES)  
 WGS. *See* Whole-genome sequencing (WGS)  
 White matter hyperintensities (WMH), 190  
 WHO. *See* World Health Organization (WHO)  
 Whole-exome sequencing (WES), 57, 61–62, 123, 394, 437  
 Whole-genome sequencing (WGS), 57, 62, 123, 184, 295–296, 303–304, 394, 437, 536  
   on DNA, 362  
 Whole-transcriptome sequencing, 373–374  
 Wiedemann–Steiner syndrome, 98f  
 Wilms tumor factor 1 (WT1), 111  
 Wilson’s disease, 480–481  
 WIP14. *See* WD repeat domain phosphoinositide-interacting protein 4 (WIP14)  
 Wiscott–Aldrich syndrome, 349  
 With-no-lysine kinases (WNK), 173  
   genes, 173–175  
   kinases, 173–174, 174f  
 WMH. *See* White matter hyperintensities (WMH)  
 WNK. *See* With-no-lysine kinases (WNK)  
 Wnt family member (WNT), 110–111  
   WNT/ $\beta$ -catenin pathway, 113  
 World Health Organization (WHO), 73, 275–276, 417  
   criteria for diabetes mellitus, 276f  
   fact sheet on obesity, 73  
 WT1. *See* Wilms tumor factor 1 (WT1)

## X

X-linked chronic granulomatous disease, 99, 380  
 X-linked Duchenne muscular dystrophy, 513  
 X-linked Emery–Dreifuss muscular dystrophy, 95  
 X-linked Hunter’s disease, 508  
 X-linked hypophosphatemia (XLH), 229  
   management of patients with, 229  
 X-linked recessive disorders, 345

X-linked SCID, 99  
 X-linked sideroblastic anemia, 336  
 X-linked single-gene disorders, 107  
 Xanthomata, 252–254  
*Xenopus* oocytes, 176–177  
 XLH. *See* X-linked hypophosphatemia (XLH)  
 Xlmr–Hypotonic Facies syndrome, 90–91  
 XLP, 409  
 46, XX DSD, 113–115  
   exposure or overproduction of androgens, 114–115  
   ovarian development, 113  
   steroid synthesis defects, 114–115  
 XY disorders of sexual differentiation, 110–113  
   disorders of androgen response, 112–113  
   disorders of androgen synthesis, 111–112  
     with adrenal dysfunction, 111  
     without adrenal dysfunction, 112  
   disorders of testicular (gonadal) development, 110–111

## Y

Yamanaka model of iPSCs generation, 509, 509f

## Z

ZFPM2, 126  
 Zinc finger homeobox 3 gene (*ZFHX3*), 195, 199  
 Zinc-finger nucleases (ZFNs), 517–518  
 Zonula adherens, 21  
 Zonula occludens, 20



# CLINICAL MOLECULAR MEDICINE

## PRINCIPLES AND PRACTICE

*Clinical Molecular Medicine: Principles and Practice* presents the latest scientific advances in molecular and cellular biology within a clinical context. The field of clinical molecular medicine is vital to the development of new and effective drugs and biological therapies and diagnostic methods. This book provides medical and biomedical students and researchers with a clear and clinically relevant understanding of the molecular basis of human disease based on the latest scientific evidence. With an increased focus on new practice concepts such as stratified, personalized, and precision medicine, this book is a valuable and much-needed resource that unites the core principles of molecular biology with the latest and most promising genomic advances in order to further the field and promise of personalized and precision clinical molecular therapeutics.

### Key Features

- Illustrates the fundamental principles and clinical applications of molecular and cellular biology
- Offers a clinically focused account of clinical and molecular heterogeneity
- Includes comprehensive coverage of various disorders, including growth and development, cardiovascular, metabolic, skin, blood, digestive, inflammatory, neuropsychiatric disorders, and many more
- Introduces the rapidly evolving practice of personalized and precision medicine

### Related Titles

- Kumar and Antonarakis, *Medical and Health Genomics*, June 2016, 358 pp., 9780124201965
- Kumar and Chadwick, *Genomics and Society: Ethical, Legal, Cultural and Socioeconomic Implications*, November 2015, 446 pp., 9780124201958



ACADEMIC PRESS

An imprint of Elsevier

[elsevier.com/books-and-journals](http://elsevier.com/books-and-journals)

