

---

## Genetic and Breeding Aspects of Lactation

---

Vijay Kumar

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/66179>

---

### Abstract

Animal breeding is the most important tool to improve milk productivity in dairy animals. Identification of better animals as parent for the next generation is the approach to generate high-producing animals. Genetic disease is the main problem in livestock sector. Complex vertebral malformation (CVM), bovine leukocyte adhesion deficiency (BLAD) syndrome, bovine factor xi deficiency, deficiency of uridine monophosphate synthase (DUMPS), and bovine citrullinemia are some of the main genetic disorders, which should be studied in Holstein and its cross. Recent developments in biotechnology have made it feasible to reveal a large number of genetic polymorphisms at the DNA level. As a result, researchers have been encouraged to use them as markers in order to evaluate genetic basis for the observed phenotypic variability. The selection of genetically superior sires is of utmost importance for any successful breed improvement program as semen of sires is disseminated in various herds under progeny testing program. The investigations on mode of genetic improvement around the world indicated that proper genetic evaluation and selection of bulls bring about more than 75% of the actual genetic improvement.

**Keywords:** molecular markers, genetic disorder, part-lactation milk yield, test-day milk yield

---

### 1. Introduction

Lactation is one of the most important economic characters of dairy animals. Genetic aspect of lactation includes those concepts, which are important for milk production in an individual animal. On the other hand, breeding aspect is important for improvement in milk production in a population of dairy animals. Milk production is largely affected by a combination of factors, namely, genetic makeup, a favorable nutritional environment and improved managerial practices. Consequently, genetic makeup of dairy animals plays a great role in the variation of milk yield and composition. Lactation is, therefore, a factor of genotype-environment interactions. Animals are subjected to constant improvement for this trait, which is based on

phenotypic observations of cattle and evaluation of the progeny and kins. These processes are both time-consuming and expensive. With the observed advances in molecular genetics, one may presume, however, that this may dramatically accelerate genetic improvement of dairy cattle, especially due to genotyping of young animals. Identification of various genetic factors involved in regulating the lactation and involution in farm animal would enable to reveal the molecular basis of milk production. This would serve as a platform to direct future research to enhance the productivity of cattle. The identification and characterization of these factors will pave the way to design appropriate approach to ameliorate the problem of lactation. The availability of fully sequenced genomes and further development of DNA sequencing techniques allowed gaining of complex genomic data from a number of animals and promises to overcome traditional problems with understanding of fragmented genomic data provided by other genomic techniques. Association studies using whole genome sequence from animals with different phenotypic performances will permit identification of genomic regions with significant effect on main economic traits. Genetic improvement includes selection of excellent animals from a population to produce higher yields in future generations. Since long back, livestock breeders have used genetic evaluations to identify best animals. Selective use of these livestock improved phenotypic measures for milk production and milk components.

## 2. Molecular markers and milk production

Recent developments in biotechnology have made it feasible to reveal a large number of genetic polymorphisms at the DNA level. As a result, researchers have been encouraged to use them as markers in order to evaluate genetic basis for the observed phenotypic variability. The distinctive genetic properties as well as methodological advantages of molecular markers make them helpful and amenable, to a greater extent, for genetic research than other genetic markers. Nowadays, molecular markers are being identified on a vast variety of genes of economic importance and are widely accepted. Association between DNA polymorphism and milk production traits has been studied for a number of genes, including prolactin; leptin; diacylglycerol acyltransferase (DGAT1); bovine leukocyte antigen (BoLA)-DRB3, growth hormone receptor gene, ATP-binding cassette, and protease inhibitor gene; proliferator-activated receptor gamma and coactivator (CoA) 1 $\alpha$  gene; growth hormone (GH) gene; signal transducer and activator of transcription (STAT)1; thyroglobulin gene;  $\beta$ -lactoglobulin gene; STAT5A; and stearoyl-CoA desaturase. Out of the different marker studies, molecular markers for prolactin, leptin, and DGAT1 show their association with milk traits.

### 2.1. Pit-1 transcription factor

Pit-1 transcription factor regulates the expression of growth hormone and prolactin in the anterior pituitary gland [1]. The bovine Pit-1 gene is located in centromeric region of chromosome 1 in bovine [2]. Bovine Pit-1 is a 291 amino acid protein. This gene is a candidate for lactation yield because of its role in regulating expression of bovine growth hormone (bGH) and the prolactin genes. Association studies have shown that Pit-1 is associated with milk production in domestic animals. Pit-1 was found to be related to birth weight weaning and average

daily gain. Quantitative trait loci (QTL) detection revealed that the region surrounding POU domain, class 1, transcription factor 1 (POU1F1) on cattle, 1q21–q22, had an effect on animal production. POU1F1 gene is also a possible candidate for growth traits. In cattle, Pit-1 was found to be associated with milk production traits. *Hinf1* polymorphism has been reported in exon 6 of the bovine Pit-1 gene by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) technique. This single nucleotide polymorphism (SNP) was found in the coding region of the bovine Pit-1 gene. Moreover, four intronic polymorphisms were also reported: two located in intron 3, one in intron 4, and the last in intron 5.

## 2.2. Growth hormone receptor (GHR)

The growth hormone plays a critical endocrine role controlling nutrient metabolism in dairy cattle. In the liver, growth hormone receptor (GHR) is dynamically regulated by lactation and energy balance. GH receptor is determined by a single gene with a length of 110 kbp, consisting of ten exons, located on chromosome 20 of cattle [3]. The gene of this receptor (GHR) is characterized by the existence of numerous exons whose transcripts undergo alternative splicing. Growth hormone activity depends on the receptor, and consequently its gene may be a potential candidate gene.

## 2.3. Insulin-like growth factor 1 (IGF-1)

Insulin-like growth factor 1 (IGF-1) is produced in the liver and body tissues. Important role for IGF-1 is to promote cell growth and division. The bovine *IGF-1* gene is located on the long arm of chromosome 5. It contains seven exons, separated by long introns. As stated, IGF-1 regulates cell growth, development, and division; it can stimulate growth in both normal and cancerous cells. Even small increases in serum levels of IGF-1 are associated with increased risk for several common cancers.

## 2.4. Diacylglycerol acyltransferase 1 (DGAT1)

DGAT1 is the main enzyme for the synthesis of triacylglycerols, the major fraction of milk fat compounds. Bovine *DGAT1* locus is found on chromosome 14. Milk is the main source of saturated fatty acids, which are a valuable source of nutrients. Recently, a dinucleotide substitution located in the exon 8 of the gene coding for acyl-CoA, diacylglycerol acyltransferase 1 (DGAT1), that modifies the amino acid order from a lysine to an alanine (Lys232Ala) in the mature protein was shown to have an immense effect on milk fat content in some cattle breeds. The frequency of polymorphism in *DGAT1* gene has been found to be very high in dairy cattle. The attention in the bovine *DGAT1* gene has increased during the previous few years. *DGAT1* is also important for the physiological processes involving triacylglycerol metabolism such as intestinal fat absorption, adipose tissue formation, and lactation.

## 2.5. Leptin (LEP)

Leptin is a 16-kDa protein secreted from white adipocytes, which is involved in regulation of feed intake, milk yield, energy balance, fertility, and immune functions in animals. Leptin is considered as a powerful biomolecule for enhancing productivity in farm ani-

mals. Its role in lactogenesis, colostrum secretion, galactopoiesis, and immunity to mastitis has made it an important candidate gene for genetic studies. Since the bovine leptin gene has been recognized on chromosome 4, several SNPs have been in the past identified in introns and exons of leptin among different breeds of cattle. Several investigations have revealed that polymorphisms were associated with lean and fat cattle [4, 5] and with the fat contents and feed intakes [6–8]. Important profitable factors for livestock production may be affected by leptin including feed conversion competence and intramuscular fat, which is considered to improve meat quality [9].

Nutrition is a main determinant of productive potential in cattle and other mammals. Therefore, responses of the productive system to changes in nutrition and metabolic status influence reproductive and economic efficiency of food-producing species in a variety of contexts. As the hormone leptin is involved in regulation of nutritional status and reproductive function, this hormone is an interesting protein to investigate during the periparturient period in dairy cattle. Leptin may help regulate ovarian development and steroidogenesis and serve as either a primary signal initiating puberty or as a permissive regulator of sexual maturation. Leptin is supposed to be necessary by the mammary growth, development, and function, when mammary fat cell leptin expression requires prolactin which then cooperates with leptin to influence mammary activity. The ruminant mammary epithelial cells also synthesize leptin during pregnancy and during established lactation. In dairy cattle, the increase in milk yield has been accompanied by a more negative energy balance during early lactation and a decrease in fertility. The polymorphism in LEP gene was associated with milk performance and increased perinatal mortality in dairy calf birth and weaning weights in beef and dairy and reproductive performance in dairy cattle. Exogenous administration of leptin in ob/ob mice leads to a significant improvement in reproductive and endocrine status as well as reduced food intake and weight loss.

## 2.6. Prolactin

The bovine prolactin gene (*bPRL*) is regarded as a good candidate gene for marker-assisted selection (MAS) for milk production parameters. Bovine prolactin spans approximately 9.4 kb and consists of five exons and four introns, which encode a mature protein with 199 amino acids first mapped it on the bovine chromosome 23. Prolactin gene disruption experiments proved its important roles in mammary gland development (mamrogenesis), lactogenesis, maintenance of milk secretion (galactopoiesis), and expression of milk protein genes [10]. It is also chiefly accountable for synthesis of lactose, lipids, and all other major components of milk [11]. Consequently, the bovine prolactin gene seems to be an outstanding candidate for linkage analysis with quantitative trait loci (QTL) affecting milk production traits. Prolactin is a polypeptide hormone and plays crucial roles in mammary gland development, initiation, and maintenance of lactation, which is synthesized in lactotroph cells of the anterior pituitary gland of vertebrate. Prolactin is a protein hormone primarily, but not solely produced by lactotroph cells of the anterior pituitary. Its role in maintenance of milk secretion is well demonstrated. As a result, the gene encoding prolactin is assumed to be one of the key links in constituting the hereditary component of milk productivity. Therefore, the bovine prolactin gene seems to be an excellent candidate for quantitative trait loci (QTL) affecting milk production traits.

### 3. Genetic disorders and animal selection

Apart from selection of bulls on the basis of milk production and reproductive abilities, some of the genetic disorders should also be taken care of. Chromosomal abnormalities in domestic animals are one of the very important causes of reproductive impairment. The impairment in reproduction leads to loss of germplasm, and there is every possibility that these chromosomal aberrations are transmitted from one generation to another. Several centric fusions, i.e., *Robertsonian* translocation, have been reported in cattle. The animals carrying this translocation are phenotypically normal but have reproductive problems. In this translocation, two nonhomologous chromosomes break at the extreme end at the short arm. These small segments are lost, and the large segments fuse at their centromeric region and produce a new large metacentric or submetacentric chromosome. The exchange of segment between two nonhomologous chromosomes is a type of structural variation called a reciprocal translocation. In this, the genetic information is neither lost nor gained rather there is a rearrangement of genetic material. The least complex way for this event to occur is for nonhomologous chromosome arms to come close to each other so that an exchange is facilitated. For this type of translocation in which the ends of chromosomes are involved, only two breaks are required. Heterozygous homologues for a reciprocal translocation undergo unorthodox synopsis during meiosis, resulting in a cross-like configuration. By whatever mechanism produced, the translocated region may nevertheless act normally with respect to homologous pairing. Thus, a translocation homozygote forms the same number of homologous pairs as the normal homozygote as long as centromeres are not lost. In several countries, there is restriction on the export and import of semen from bulls without normal karyotype. In the United Kingdom, all the bulls used for artificial insemination are subjected to routine screening for cytogenetic abnormality.

Selection is based on genetic markers and also universally checking up of the breeding bulls for heritable diseases. Many heritable mutations and diseases have been traced to autosomal chromosomes of different cattle breeds; most of these mutations are lethal or bring about deficiencies in the defense mechanism if they occur in the homozygous recessive manner. Moreover, due to coverage of wild-type allele on mutant allele in heterozygotes, carrier animals may not show any phenotypic abnormality but are most likely to transmit the mutant alleles to the next generation. The same problem may assume catastrophic significance if the carrier is a bull intended to be used in artificial breeding programs.

#### 3.1. Complex vertebral malformation

The complex vertebral malformation (CVM) is a recessively inherited congenital disorder leading to frequent abortion of fetuses or vertebral anomalies and prenatal death [12–14]. It is characterized by growth retardation and bilateral flexure of the carpal and metacarpophalangeal joints along with rotation of the digits. In stillborn, aborted, and preterm calves, CVM has been characterized by shortened cervical and thoracic regions of the vertebral column and symmetric arthrogryposis. Multiple hemivertebrae, scoliosis, and synostosis and fused and misshaped vertebral column have also been described. The syndrome was first

discovered in the Danish Holstein population in 1999, but, shortly thereafter, its occurrence was reported in the Netherlands, in the United States, in the United Kingdom, and in Japan.

Studies of Danish Holstein showed that the extent of fetal mortality was approximately 77% prior to gestation day 260. This is reflected in a significantly reduced ratio of CVM-affected calves in breeding studies. Through increased fetal mortality, the status of the fetus has a significant effect on calving interval and involuntary culling of cows. The symptoms of the defect have not been observed in carriers of CVM. No productive and reproductive differences between carrier and noncarrier animals have been reported. The only difference which is very important was increase in the rate of intrauterine mortality. The risk of return to service was higher in carrier animals.

When studying the postpartum commencement of ovulation—the carry-over of ovarian cycles after the birth of one calf—the researchers recognized that there was no major difference in the rate at which resumption occurs in carriers as compared to Holstein cows that were not carriers for CVM. However, CVM carrier cows were recognized to have drastically lower conception rates than control cows, with the gap between calves observed at 463 days in carriers than 399 days for control cows, telling the possibility of heightened death of embryos in CVM carrier cows.

### **3.2. Bovine leukocyte adhesion deficiency syndrome**

Bovine leukocyte adhesion deficiency (BLAD) syndrome is an autosomal recessive hereditary disease affecting young Holstein calves characterized by recurrent bacterial infections, progressive periodontitis, ulcers of oral mucosa, and impaired inflammatory responses [15, 16]. These clinical findings are associated with impaired neutrophil functions such as markedly decreased adherence, chemotaxis, and phagocytosis. In 1990, a lack of  $\beta_2$  integrin molecules expressed on the leukocytes from affected animals was found in a calf with granulocytopenia syndrome, and this disease was termed bovine leukocyte adhesion deficiency, which was considered to be analogous to human leukocyte adhesion deficiency (LAD). BLAD causes immune deficiency in the early days of life. The affected animal is prone to all infectious agents which can cause death. BLAD was first documented in Holstein-Friesian cattle, and no study has reported the incidence of this disorder in other breeds. The affected cattle with BLAD was linked to common ancestral sires that had been documented to be carriers.

### **3.3. Deficiency of uridine monophosphate synthase**

Deficiency of uridine monophosphate synthase (DUMPS) is a genetic disorder which interferes with pyrimidine biosynthesis. Uridine monophosphate synthase (UMPS) has a key role on the pyrimidine nucleotide synthesis, which is essential for normal growth and development for several ruminant and nonruminant species [17]. Inactivation of this enzyme is caused by an autosomal recessive heredity mutation, which takes place in the gene of uridine monophosphate synthase. The mutation (cytosine to thymine) leads to the loss of the restriction site of *Ava*I site at nucleotide position 405 of the gene. This disorder is named as DUMPS in the Holstein cattle and characterized by lowered blood activity of enzyme UMPS [17]. DUMPS

leads to embryonic death in the early stage of pregnancy. So some serious reproductive problems take place in dairy herds.

### 3.4. Factor XI deficiency syndrome

Factor XI (FXI) deficiency is an autosomal recessive disorder, with partial deficiency of FXI coagulant activity in heterozygotes and considerable deficiency in homozygotes [18]. One of the protein factors concerned in blood coagulation is a serine protease—factor XI. It is produced in the liver as a zymogen, and after alteration to a proteolytic enzyme. FXI deficiency has been recognized in humans and cattle. FXI gene is located on chromosome number 4 in human and on chromosome number 27 in cattle. FXI gene encodes coagulation factor XI of the blood coagulation cascade, and it is one of more than a dozen proteins involved in blood clotting.

In cattle, FXI-deficient animals may be asymptomatic or display several symptoms, like prolonged bleeding, anemia, greater prevalence of repeat breeding, or even lower resistance to pneumonia, mastitis, and metritis. *FXI* deficiency has been recognized in a number of species of mammals, including humans, dogs, and cattle. *FXI* deficiency possibly will result in delayed bleeding and anemia. Persistent bleeding from the umbilical cord is occasionally seen in affected calves. Delayed emission of blood following dehorning or castration may also be observed. Affected cows repeatedly have pink-colored colostrum.

### 3.5. Bovine citrullinemia

Bovine citrullinemia is a genetic disorder, which has been reported in Holstein cattle [19]. It has been established that bovine citrullinemia is a consequence of a deficiency of argininosuccinate synthetase (ASS), one of the enzymes of the urea cycle. A shortage of the urea cycle enzyme consequences in a lethal neurological disease in newly born calves. The urea cycle entails a series of biochemical steps in which waste product of protein metabolism is removed from the blood in the form of urea. The deficiency of ASS occurs when a calf inherits a copy of the mutant gene encoding for ASS from each parent. The gene is located on chromosome number 11 (BTA11). The mutation is caused by a transition of cytosine to thymine at codon 86 within exon 5 in the gene coding for ASS leading to impaired enzyme, which cannot participate in urea cycle. Calves affected with the disorder appear normal immediately after birth. However, by the second day of birth, they become depressed. By the third day, they are seen wandering or standing with their head pressed against a wall. Death usually takes place within 12 hours of start of these clinical signs. The clinical signs of citrullinemia are believed to be as a consequence of accumulation of ammonia in the brain of the affected calves.

## 4. Factors affecting milk production

The success of dairy industry is much dependent on productivity and efficient reproduction performance of animals. Genetic progress may be done by selection. The nongenetic

factors such as management, housing, feed, season, etc. also control production. The genetic improvement of growth of dairy animals is of great importance in the large ruminant industry.

#### **4.1. Dry period**

Dry period is an important economic trait which has the direct effect on lifetime milk production. There is a lot of variability observed in the dry period. It is the period from the date of drying to the next calving. Animals should be given proper rest, when they are in pregnancy, before the next calving to compensate for growth of fetus.

#### **4.2. Growth**

The genetic improvement of growth is of great importance in the large ruminant industry. The growth rate of a heifer calf until it matures to a cow is also a significant trait since it characterizes the adaptability and economic appropriateness of the animal. It is anticipated that animals growing faster may also begin physiological performance of reproduction and milk production earlier.

#### **4.3. Age at first calving**

Age at calving is an important trait because lower age at first calving (AFC) leads to shorter generation interval and hence increases genetic gain. AFC has a considerable influence on the total expenditure of raising dairy replacements. Older calving heifers are being more expensive to raise than younger calving heifers. Moreover, reducing AFC can also increase the profitability by increasing lifetime milk production and milk production per year of herd life. The most important drawback of reducing AFC is that it is commonly associated with a reduction in the first lactation milk yield. Regardless of this decline in the first lactation milk yield, production per year of herd life is characteristically increased by reduced AFC. In addition, while the first lactation may be affected by AFC, future lactations are absolutely not. Furthermore, stayability and health of cows are not affected by reduced AFC as long as the first calf heifers freshen at a sufficient weight.

#### **4.4. Service period**

Service period is the duration between date of calving and date of successful conception. The finest service period facilitates the animal to recover from the stress of calving and also for reproductive organs back to its usual shape. Usually for cattle the best service period is 60–90 days.

#### **4.5. Calving interval**

The interval between two successive calvings is referred to as calving interval (CI) which consists of two components—service period and gestation period. Among the traits of dairy animals, calving interval is the most important trait as it determines the number of lactations possible in the lifetime of an animal, thereby influencing its production of an

animal. It is generally held that shorter calving intervals favor a large yield for the lifetime of the cow. That frequent calving should give maximum lifetime yield seems in harmony with the known facts of the physiology of reproduction and lactation. The natural need of milk develops with birth of the young. The need increases with growth of the young so long as milk is the only food. As other food is taken, the need for milk decreases and finally disappears. In milch animals, where we have adequate information as to the amount of milk produced day by day, it is found that the rate of milk secretion parallels roughly the rising and declining needs of the young. That is, the rate of milk secretion is most rapid in the cow shortly after calving and slows up quite markedly and continuously thereafter. It would seem that frequent calving by giving more high points in the lifetime lactation curve should result in a greater total yield than less frequent calving. Another way of putting the matter is to relate the length of the calving interval to the average yield of milk per day over the calving interval. This relation may be derived in relative terms from the known characteristics of the lactation curve. The indication is that the highest lifetime yield should be obtained from calving intervals of less than a year rather than from calving intervals of more than a year.

#### **4.6. Days to first service or waiting period**

Waiting period (WP) or days to first service (DFS) is one of the important reproduction traits. WP or DFS is the initial phase of lactation during which no inseminations occur. The lower the WP or DFS shows, the higher the breeding efficiency in dairy animals. Reducing WP is tempting because of associated reductions in calving interval. Getting cows pregnant is perhaps one of the major challenges faced by the dairy farmers. Cows that do not become pregnant at a rational interval after calving will remain in the herd, reduced overall milk yield, and if they do become pregnant, they likely to develop post-parturient diseases in the subsequent lactation. Conversely, on the other hand, cows that become pregnant too early after calving would be dried off too early in lactation when they are still producing a reasonable amount of milk.

The voluntary waiting period (VWP) is a time during the early lactation in which cows are purposely not inseminated even though they show estrus signs, to allow for optimum uterine involution. The VWP is fairly inconsistent across dairy herds. Decision over the optimal VWP should be made according to individual situations and after careful analysis of probable gains and losses. The suggestion of extending the VWP is based on the reality that during the early lactation cows are recovering from a condition of negative energy balance. Therefore, extending the VWP may allow cows to recover from a metabolic state that is deleterious to reproductive efficiency and display estrus a few times prior to the first insemination, which has been correlated with improved fertility.

#### **4.7. Daughter pregnancy rate**

A genetic evaluation tool for reproduction exists in the form of daughter pregnancy rate (DPR). DPR is calculated from days open and is directly related to the proportion of females eligible to become pregnant in a 21-day period that actually become pregnant (i.e., the 21-day pregnancy rate).

DPR forecasts genetic improvement (or deterioration) in pregnancy rate for future daughters of a bull compared to a bull that is expected to produce no change. Better pregnancy rates decrease semen usage and can be an essential part of efforts to reduce service period. Workers also comment that genetics do not control fertility adequately to justify selection. Heritability of DPR is low—only about 4%—so producers should not expect dramatic improvement in fertility from selection. However, genetic control of fertility is real. DPR needs to be included with all the other traits of economic significance in dairy cattle breeding.

## 5. New breeding approaches to improve dairy animals

The selection of genetically superior sires is of utmost importance for any successful breed improvement program as semen of sires is disseminated in various herds under progeny testing program. The investigations on mode of genetic improvement around the world indicated that proper genetic evaluation and selection of bulls bring about more than 75% of the actual genetic improvement. The accuracy of estimating the breeding value of an animal is the major factor that affects the genetic progress from selection. The success of a breeding program depends on how early and how accurately young bulls can be evaluated at a minimum possible cost.

Animal breeding is the most important tool by which one can think about improving livestock sector in comparatively faster rate. Animal breeding is all about selecting best animals from population which are used as parent for the next generation. Milk yield is the single most important economic trait determining economic returns from the dairy animals. In animal breeding selection of animals is almost considered as selection of bulls since the progeny-producing ability of bull is many folds more than a cow and progeny testing is the most acceptable technique for the selection of bulls.

### 5.1. Part-lactation milk yield for genetic evaluation

The first important advancement was the use of part-lactation milk yield for genetic evaluation of cattle and buffalo bulls. The genetic and phenotypic correlation between 305-day and part-lactation monthly and cumulative monthly milk yields in cattle reported that the genetic correlation of monthly yields with 300-day yield was high up to the fifth month and thereafter the estimates declined. The phenotypic correlations among these traits were high and statistically highly significant. There is significant correlation between 100-day and 300-day milk yields, and selection on the basis of 100-day production was sufficiently accurate to select the animal for 300-day production in Gir cattle. In Sahiwal cattle selection on the basis of individual third month milk yield and cumulative 90-day milk yield was at least 89 and 79% as effective as on the basis of first lactation 305-day milk yield. Definitely, these findings will be very helpful in reducing duration for evaluating bulls as only 4–5 months milk records are sufficient enough to evaluate bulls. Hence, the application of part-lactation records of daughters will be helpful in ranking the sires at younger age resulting in reduced generation interval

and increased intensity of selection credited to the availability of more numbers of records on daughters even having part-lactation records.

### 5.2. Test-day milk yield for genetic evaluation

The second important advancement is the use of test-day milk yield for genetic evaluation of cattle and buffalo bulls. The use of test-day record is easier and advance than part-lactation record as in this case only one or two particular day (test day) record is required instead of taking all-day records up to a certain part of lactation. Genetic correlations between observed milk yield and predicted milk yield under various test-day sampling schemes (fortnightly, monthly, and bimonthly) were very high being greater than 0.95 for Karan Swiss cattle and to be nearly unity in Karan Fries cattle. There are higher phenotypic correlations among test-day milk yields ranging from 0.55 to 0.93 in Murrah buffaloes. The genetic correlations among test-day milk yields and 305-day milk yield were close to unity and phenotypic correlations among test-day milk yields and 305-day milk yield were quite high and positive (0.60–0.99) in Murrah buffaloes. So this technique was also explored at research level and found that test-day records could also be used for fast progress in livestock improvement.

At the present time, test-day model is thus used instead of 305-day model for the genetic evaluation of dairy cattle. Test-day model is the statistical method that thinks about all genetic and environmental effects straight on a test-day basis. The test-day models have been recommended as the method of choice for the study of milk yield traits in order to maximize the use of all accessible information. This method becomes significant with smaller herd size and without well-established milk recording systems.

## 6. Conclusion and future perspectives

A variety of molecular markers for improving milk production and their association with disease and productive and reproductive traits have proved to be beneficial to the livestock breeders. They can also be used economically in breeding and management decisions. The application of diverse molecular markers in determining the vulnerability to economically significant diseases proves to be helpful to minimize loss of animals and their productivity. Advancement of technologies to evaluate polymorphisms provides an idea about the development in the field of genetics to improve cattle health and production. The genetic disease causes heavy losses because of poor animal performance; structural unsoundness reduces the production and reproductive potential of the animal. The massive spread of genetic defects like CVM and BLAD in recent years is caused by the extensive use of elite carrier sires. Artificial insemination hastens the spread of detrimental recessives worldwide. The new technologies of molecular genetics facilitate to find the cause at gene level. This makes it possible to identify heterozygous animals and to manage the genetic health of the cattle population. The use of molecular methods into animal research promises quick progress in animal potential. If no cases are found in routine screening even then, extensive screening programs for

identification of genetic disorders appear necessary to guarantee the utilization of bulls free from genetic disorders for artificial insemination programs.

## Author details

Vijay Kumar

Address all correspondence to: [jaisvet@gmail.com](mailto:jaisvet@gmail.com)

Department of Animal Genetics and Breeding, U.P. Pt. Deen Dayal Upadhyaya Pashu Chikitsa Vigyan Vishwavidyalaya Evam Go Anusandhan Sansthan, Mathura, Uttar Pradesh, India

## References

- [1] Lin C, Lin SC, Chang CP, Rosenfeld MG. Pit-1-dependent expression of the receptor for growth hormone releasing factor mediates pituitary cell growth. *Nature*. 1992;360:765–768. DOI: 10.1038/360765a0
- [2] Moody DE, Pomp D, Barendse W. Restriction fragment length polymorphism in amplification products of the bovine PIT1 gene and assignment of PIT1 to bovine chromosome 1. *Animal Genetics*. 1995;26:45–47. DOI: 10.1111/j.1365-2052.1995.tb02620.x
- [3] Moody DE, Pomp D, Barendse W, Womack JE. Assignment of the growth hormone receptor gene to bovine chromosome 20 using linkage analysis and somatic cell mapping. *Animal Genetics*. 1995;26:341–343. DOI: 10.1111/j.1365-2052.1995.tb02671.x
- [4] Stone RT, Kappes SM, Beattie CW. The bovine homolog of the obese gene maps to chromosome 4. *Mammalian Genome*. 1996;7:399–400. DOI: 10.1007/s003359900119
- [5] Fitzsimmons CJ, Schmutz SM, Bergen RD, McKinnon JJ. A potential association between the BM 1500 microsatellite and fat deposition in beef cattle. *Mammalian Genome*. 1998;9:432–434.
- [6] Konfortov BA, Licence VE, Miller JR. Re-sequencing of DNA from a diverse panel of cattle reveals a high level of polymorphism in both intron and exon. *Mammalian Genome*. 1999;10:1142–1145. DOI: 10.1007/s003359901180
- [7] Buchanan FC, Fitzsimmons CJ, Van Kessel AG, Thue TD, Winkelman-Sim DC, Schmutz SM. Association of a missense mutation in the bovine leptin gene with carcass fat content and leptin mRNA levels. *Genetics Selection Evolution*. 2002;34:105–116. DOI: 10.1051/gse:2001006
- [8] Lagonigro R, Wiener P, Pilla F, Woolliams JA, Williams JL. A new mutation in the coding region of the bovine leptin gene associated with feed intake. *Animal Genetics*. 2003;34:371–374. DOI: 10.1046/j.1365-2052.2003.01028.x
- [9] Wheeler TL, Cundiff LV, Koch RM. Effect of marbling degree on beef palatability in *Bos taurus* and *Bos indicus* cattle. *Journal of Animal Science*. 1994;72:3145–3151.

- [10] Horseman ND, Zhao W, Montecino-Rodriguez E, Tanaka M, Nakashima K, Engle SJ. Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. *EMBO Journal*. 1997;16:6926–6935. DOI: 10.1093/emboj/16.23.6926
- [11] Le Provost E, Leroux C, Martin P, Gafe P, Dijane J. Prolactin gene expression in ovine and caprine mammary gland. *Neuroendocrinology*. 1994;60:305–313.
- [12] Agerholm JS, Andersen O, Almskou MB, Bendixen C, Arnbjerg J, Aamand G, Nielsen US, Panitz F, Petersen AH. Evaluation of the inheritance of complex vertebral malformation syndrome by breeding studies. *Acta Veterinaria Scandinavica*. 2004;45(3):133–137. DOI: 10.1186/1751-0147-45-133
- [13] Agerholm JS, Bendixen C, Andersen O, Ambjerg J. Complex vertebral malformation in Holstein calves. *Journal of Veterinary Diagnostic Investigation*. 2001;13:283–289. DOI: 10.1177/104063870101300401
- [14] Nielsen US, Aamand GP, Andersen O, Bendixen C, Nielsen VH, Agerholm JS. Effects of complex vertebral malformation on fertility traits in Holstein cattle. *Livestock Production Science*. 2003;79(2):233–238. DOI: 10.1016/S0301-6226(02)00170-7
- [15] Kehrli ME, Ackermann MR, Shuster DE, Maaten VD, Schmalstieg FC, Anderson DC, Hughes BJ. Bovine leukocyte adhesion deficiency.  $\beta_2$  integrin deficiency in young Holstein cattle. *American Journal of Pathology*. 1990;140:1489–1492.
- [16] Nagahata H, Noda H, Takahashi K, Kurosawa T, Sonoda M. A suspected case of neutrophil dysfunction in a Holstein heifer. *Japanese Journal of Veterinary Science*. 1987;49(6):1165–1167.
- [17] Healy MH, Shanks RD. Performance of females heterozygous for deficiency of uridine monophosphate synthase. *Journal of Dairy Science*. 1987;70(5):945–951.
- [18] Gentry PA, Ross ML. Coagulation factor XI deficiency in Holstein cattle: expression and distribution of factor XI activity. *Canadian Journal of Veterinary Research*. 1994;58(4):242–247.
- [19] Fesus L, Anton I, Zsolnai A. Marker assisted selection in livestock. DUMPS, weaver-diseases and citrullinaemia in cattle population. *Allatt-es-Takarm*. 1999;48(2):193–203.



# Genetics in Domestic Animal Reproduction

---

Sven Budik

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/67132>

---

## Abstract

Reproduction and early development in domestic animals are biological processes in which complex genetic events like formation of the germ cells, meiosis, syngamy, zygote formation, cleavage, midblastula transition, dose compensation of sex chromosomes, genetic imprinting, and multiple cell differentiation take place. Many of these processes have great impact in veterinary reproductive medicine and are influenced by the assisted reproductive techniques applied. Altered environmental and metabolic conditions caused by intensive livestock farming influence reproductive success by altered gene expression via epigenetic changes.

**Keywords:** reproductive genetics, developmental genetics, domestic animals

---

## 1. Genetics during germ cell production

Genetics during germ cell production: Mammalian oocytes and sperms originate from the primordial germ cells (PGCs) located in the embryonic mesoderm. Their development is initiated by signals from the extra embryonic ectoderm and the visceral endoderm. In mammals the primordial germ cells invade the genital ridge where they proliferate by mitosis and give rise to either oogonia or spermatogonia. Migration, proliferation, and colonization of PGCs to the developing gonads are controlled by many factors and depend as well on the interaction of PGCs and their surrounding somatic cells. Around the period of PGC migration into the genital ridges sex determination starts. Absence of the expression of the Y-linked gene "sex determining region of Y" (SRY, a transcription factor of the high mobility group box family) leads to female differentiation, therefore the gonads develop into ovaries [1]. Presence of SRY gene expression leads to differentiation of the gonads into testes. As soon as PGCs are formed, the initially bi-potential gonad will continue its differentiation mostly under the influence of somatic cell-derived transcription factors. In female animals after colonization

of the gonad, PGCs will undergo a phase of mitotic proliferation leading to the formation of germ cell nests. Following this event, around birth, the germ cell nests break down and mitotic divisions stop. Germ cells initiate meiosis, become primary oocytes which are surrounded by somatic cells leading to the formation of primordial follicles which continue the female program of development.

Meiosis can be divided into two divisions: meiotic division I and II. The prophase of meiotic division I is extended and can be divided into the following parts: leptotene, zygotene, pachytene, diplotene, and diakinesis. During leptotene, progressive condensation of the chromosomes takes place and the telomeres become associated with the inner nuclear membrane. In zygotene, the homologous chromosomes pair due to the formation of the synaptonemal complex in a zipper-like way. The pachytene is characterized by further condensation of the chromosomes and by occurrence of cross-over events leading to intrachromosomal recombination. In the course of the diplotene, the chromosomes separate again except small anchoring regions the so-named chiasmata. Oocytes progressed to the diplotene stage enter into a prolonged resting phase called dictyotene. Oocytes remain at the dictyate stage of meiosis I throughout oogenesis, until luteinizing hormone (LH) induces final oocyte maturation to metaphase II [2]. Cell division during oocyte meiosis is asymmetrically leading to one oocyte containing most of the cytoplasm and three polar bodies. Formation of the second polar body is a requisite of oocyte maturity.

Spermatogenesis, the process that gives rise to fertile sperm, occurs in a similar manner in many animal species, including domestic mammals and birds. Male germ cell development starts with emergence of primordial germ cells (PGCs), which migrate and associate with somatic cells that form the testes. Within the gonad, PGCs localize to regions that are competent to serve as stem cell (SC) niches, where they develop into spermatogonial SCs (SSCs). SCs divide to form new SCs and spermatogonial daughter cells. The daughter cells proliferate, undergoing sequential rounds of mitosis and incomplete cytokinesis to form syncytial groups of spermatogonia. Following a series of synchronous mitotic amplification divisions, spermatogonia differentiate into spermatocytes, which then undergo meiosis. As in the earlier mitotic divisions, meiotic cytokinesis is incomplete, resulting in syncytia of interconnected haploid spermatids. Postmeiotic spermatids undergo an extensive period of differentiation, or spermiogenesis, in which they form sperm-specific organelles, including the sperm head, acrosome, basal body, specialized mitochondria, and flagellum. At the end of spermiogenesis, mature sperm are separated from each other by removal of the remnants of incomplete cytokinesis, released from the gonad, and stored prior to delivery by ejaculation.

In both, oogenesis and spermatogenesis meiosis leads to intrachromosomal recombination (cross-over occurring in prophase of meiosis I) and in interchromosomal recombination (random mixture of maternal and paternal chromosomes during both meiotic divisions), reducing the diploid chromosome complement to a haploid level. As a result of oogenesis, one haploid oocyte and three haploid polar bodies develop. In spermatogenesis four haploid sperm come into being. In female eutherian mammals meiosis is continued by luteal hormone stimulation and finished during fertilization. In male eutherian mammals meiosis takes place all the time.

## 2. Genetics during fertilization

During their active and passive migration in the femal genital tract, the sperms obtain their fertilization capacity by capacitation and hyper activation two processes which seem to proceed largely independent. The glycoprotein composition changes during capacitation of the cell membrane. In domestic mammalian species fertilization takes place in the ampulla of the oviduct where the metaphase II oocyte surrounded by the acellular zona pellucida and the cellular cumulus (corona radiata) is present. The sperm leading to fertilization has to fulfill its acrosomal reaction. This is mediated by partial fusion of the cell membrane with the outer acrosomal membrane leading to the release of hyaluronidase. This enzyme enables the sperm to cross the corona radiata. At the zona pellucida binding to species specific zona proteins is necessary (species specificity). After detachment of the cell and outer acrosomal membrane the inner acrosomal membrane rises to the top delivering the protease Acrosin which enables the penetration of the zona pellucida. Now the sperm enters to the perivitelline space and the inner acrosomal membrane and the cell membrane of the oocyte fuse. In natural fertilization, only the head of the sperm enters the cytoplasm of the oocyte meaning that no mitochondria from the sperm can participate (maternal inheritance). After entrance of the first sperm, depolarization of the cell membrane, degranulation of the corticalis granula, shrinkage of the ooplasm, and zona hardening takes place in order to avoid multiple fertilizations. Fertilization by more than one sperm would lead to aneuploid ( $n = 3, 4, 5\dots$ ) chromosome complement. These embryos stop development before reaching the blastocyst stage. After entrance of the sperm the oocyte finishes the second meiotic division. Female and male pronuclei form and fuse to form a diploid nucleus in a process named syngamy leading to the zygote entering now into cleavage.

## 3. Genetics of eutherian sex determination and dose compensation

Sex determination in eutherian mammals is mediated by the chromosomes X and Y. Absence of the Y chromosome (SRY gene) leads to female development. Male genital development is mediated by production of the anti-Müllerian hormone (AMH) produced by the Sertoli cells of the embryonal testes. This peptide hormone, belonging to the transforming growth factors (TGF) family, leads to the regression of the Müllerian duct from which in the female the oviduct, the uterus, and the vagina are formed.

Aberrant sex chromosome numbers can arrive by false segregation during meiosis leading to multiple X or Y chromosome numbers. If one or more Y chromosomes are present the sex is male, if no Y chromosome is present the sex is female. Additional X chromosomes will be inactivated to a large extend (only a few genes escape from this inactivation) according to the Lyon hypothesis. This inactivation is mediated by the noncoding X inactive-specific transcript (XIST) RNA. The corresponding gene is located at the X Inactivation Center (XIC) at the X chromosome. Coating of the X chromosome to be inactivated by the XIST RNA leads to the deactivation of most of the genes: Only the XIST gene itself and about one quarter of the genes present mainly in the pseudo autosomal regions of the X chromosome escape

from this inactivation. Before inactivation starts small quantities of XIST RNA are produced by both X chromosomes. After inactivation XIST is expressed from the inactivated X chromosome, whereas XIST expression from the active X chromosome ceases. Cytologically the inactivated X chromosome can be visualized as Barr body. The XO syndrome named Turner syndrome is sometimes observed in mares which are not fertile. In cats, the coat color gene is located at the X chromosome leading to the absence of tortoiseshell and calico coats in the X0 (Turner syndrome) female cat. In male cats, this coloring is only possible in the rare case of the Klinefelter syndrome [3]. These animals typically have an extra X chromosome (XXY) and their cells undergo an X-inactivation process like that in females. Additional Y chromosomes are described in humans as a result of a nondisjunction in meiosis II. This chromosomal aberration seems to have less influence since only a few genes are transcribed from the Y chromosome. Nevertheless, in humans, a higher testosterone concentration was observed in chromosome pair in their pseudoautosomal region(s) (PAR). In the horse this region was very helpful in mapping of the euchromatic region of the equine Y- chromosome. These investigations were very helpful for the prediction of male fertility in horses and are therefore considered in horse breeding [4, 5].

Rarely as a consequence of a translocation, the SRY region containing the testes determining factor (TDF) to the X chromosome during meiosis a male cytological having two X chromosomes can be created. These XX men are infertile since they lack other genes necessary for the production of mobile spermatozoa [6].

Before fertilization the XIST gene in the ripe oocyte is inactive, meaning that the X chromosome is active. Fertilization by a male (Y) sperm does not change this status. If a femal (X) sperm enters the oocyte one active XIST gene is now present producing XIST RNA which inactivates also the second X chromosome. This status is not stable since for maintenance of inactivation a protein named EED (polycomb protein EED, responsible for maintaining the transcriptional repressive state of genes over successive cell generations) also coded at the X chromosome is necessary [7, 8]. Therefore, both X chromosomes and also the XIST genes become active again at the morula stage leading to random inactivation of either the maternal or paternal X chromosome. That means that female eutherian mammals are mosaics in relation to maternal or paternal X inactivation, a phenomenon which becomes evident in X coded genes like coat color of the domestic cat.

#### 4. Genomic imprinting

Genomic imprinting became evident by nuclear transplantation experiments at the pronuclear stage in mice [9]. Enucleated oocytes reconstructed with either only two femal or two male pronuclei exhibited different developmental potential concerning the embryo proper and the extraembryonic tissue. Further investigations revealed the nature of this phenomenon: In imprinted genes the female and male promotor regions have different methylation status leading to different gene expression in the early embryo. In humans and mice about 80 imprinted genes are known many of which are involved in embryonic and placental growth and development. Sometimes hybrid offspring of two different species may exhibit unusual

growth due to the novel combination of imprinted genes. This becomes especially evident in reciprocal Lion-Tiger crosses [10]. The majority of imprinted genes are found in clusters, the so-called imprinted domains, suggesting a high degree of coordinated regulation. Imprinting is a dynamic process: imprints can be generated and removed at different time points of development from one generation to the other. In the germ line the imprints are erased and then reestablished according to the sex of the individual. In the developing sperms (during spermatogenesis), paternal imprints are established, whereas in developing oocytes (during oogenesis) a maternal imprints are established. In mammals genomic imprinting is present in therian mammals (marsupials and placental mammals) [11, 12].

## 5. Genetics during preimplantative eutherian embryo development

The eutherian oocyte is oligolecithal. After fertilization, the female and male pronuclei fuse to establish the diploid zygote nucleus. The eutherian zygote undergoes now rotational holoblastic cleavage. These fast cell divisions consist only in synthesis phase and mitosis and subdivide the former ooplasm onto the blastomeres. Zygotic transcription starts dependent on the species at the two-, four-, or eight-cell stage. This process is named midblastula transition (MBT). So far the early embryo translates only maternally inherited mRNAs just present in the oocyte. So far the zygotic chromatin is hypo-acetylated and methylated, which means that most of the genes are repressed in a heterochromatic state. Now embryo starts to transcribe its own DNA and its cells become motile and the cell divisions loose synchrony. At the eight-cell stage most of the blastomeres become polarized and develop tight junctions with the other blastomeres. At the 16-cell state, the embryo is named morula. The outer cells become bound tightly together by the formation of cell junctions (compact morula). This process initiates the differentiation of two distinct cell populations: outside cells characterized by polarity and undifferentiated inside cells. The outer cells develop into trophoblast cells, which are epithelial cells connected by tight junctions in order to seal the extracellular passage. They express  $\text{Na}^+/\text{K}^+$  ATPases and aquaporins at the apical and basal membranes in order to exchange sodium in from the outside by cellular potassium. This leads to electrochemical gradient which is equalized by water influx from the outside leading to blastocoel formation. The embryo is now called a blastocyst. At least in the horse expansion of the blastocyst seems to be under hormonal control [13, 14]. After implantation the trophoblast cells will develop into the fetal part of the placenta. The undifferentiated inner cells are displaced by the water influx to one side of the cavity to form the inner cell mass (ICM). It will develop into the embryo proper and some extraembryonic membranes.

So far the blastocyst like all previous embryonic stages are still surrounded by the acellular zona pellucida since the size of the embryos increased only marginally. Now a rapid expansion mainly due to fluid accumulation takes place and therefore the embryo hatches (human, mouse, pig, cow, sheep, goat, camels) or the zona pellucida is replaced by an embryonic capsule participating in enlargement. Hatched embryos show early placenta formation whereas encapsulated embryos exhibit prolonged preimplantative phase. At that time in most species the maternal recognition of pregnancy (MRP) takes place ensuring the maternal changes necessary for further development.

## 6. Genetics during mammalian placentation and fetal development

Mammalian placentation can be rated according to the implantation depth and surface expansion. Concerning implantation depth the invasiveness of the chorion is the crucial factor: epitheliochorial (pig, horse, ruminants), endotheliochorial (dog, cat), or hemochorial (human) describe implantational depth. Surface expansion is also influenced by placental capacities and number of fetuses: diffuse, cotyledonaria, zonaria, and discoidal. Implantational depth influences the quality of exchange of substances: The human hemochorial placenta enables the transfer of maternal antibodies into the fetal blood which is very important during the first weeks after birth. From the fetal face also a transfer of fetal cells into the maternal blood stream is probable. This microchimerism can persist even for some years and might be responsible for occurrence of autoimmune diseases in women [15]. In carnivoral placentation, such a microchimerism seems to be probable but was not investigated so far. The circumstances concerning the antibodies seem to be similar like in humans: "In carnivora and rodentia the blood of the newborn has roughly the same globin content as the maternal blood, whereas in those animals in which the placental barrier is relatively thick, such as the cow and pig, the fetus receives few if any antibodies before it is born" [16]. In horses despite their epitheliochorial placentation a microchimerism was proven recently by detection of SRY by means of digital PCR [17]. One of the most probable explanations for this microchimerism seems to consist in the migration of the chorionic girdle cells forming the endometrial cups.

Another form of microchimerism often arrives in cattle during twin pregnancies with different fetal sex named Freemartin syndrome: In most bovine twin pregnancies anastomoses form between the fetal circulations leading to exchange of male hormones (anti-Müllerian hormone, testosterone). Depending on the time of anastomose formation during pregnancy the female fetus becomes masculinized resulting in intersexuality and infertility. In other domestic ruminantia the occurrence of freemartin syndrome is very rare [18].

## 7. Genetic alterations caused by assisted reproduction techniques

During *in vitro* maturation (IVM) and *in vitro* fertilization (IVF) the gamets respectively the early embryo are exposed to artificial environments which can differ in both directions from the natural conditions. Absence of essential substances often leads to the termination of development, whereas supernutrition can result in epigenetic changes which manifest later in pregnancy and during birth. One of the most frequent abnormalities connected with IVF is the large offspring syndrome in cattle and sheep resulting in oversized fetuses [19].

Somatic nuclear transplantation of domestic animals can cause incomplete epigenetic reprogramming of the somatic nucleus leading to incapacity of blastocyst formation. Therefore, blastocyst rates are relatively low around 2–3%. High rates of embryonic loss, abortions due to placental aberrations, and high postnatal mortality were observed in cloned animals. Directed epigenetic changes during somatic cell culture can ameliorate the blastocyst rate obtained (up to 10%) and reduce the problems during pregnancy and birth [20].

During intracytoplasmic sperm injection (ICSI) a mature oocyte is injected with one sperm or a haploid progenitor of the sperm. This leads to omission of the natural selection of sperms in the femal genital tract and sometimes to fertilization with a sperm that under natural conditions never would be able to fertilize an oocyte. Using ICSI no acrosomal reaction and hyperactivation is necessary. Even dead sperms can be used if the nucleus is intact. Different from natural fertilization the whole sperm is injected into the oocyte leading to a mixture of mitochondria. There seems to be a mechanism to favor the mitochondria originally contributed by the oocyte (also in cloned animals). Nevertheless also coexistence and the reverse are possible [21]. In humans oocytes from aged women can be ameliorated by cytoplasma donation from an oocyte of a young women leading to a mixture of different mitochondria similar to the situation after nuclear transplantation.

In all these techniques applied in animal-assisted reproduction, the ability of blastocyst formation is the critical point to rate the success of the method. In human-assisted reproduction, earlier stages are transferred in multiple numbers often leading to twin and multiple pregnancies.

Different other techniques like oocyte transfer (OT) gamete intrafollopiian transfer (GIFT), and zygote intrafollopiian transfer (ZIFT) have been developed. Depending on the time of incubation of the gametes and the abandonment of embryo culture these techniques have reduced potential for epigenetic alterations.

PCR techniques using genetic material from blastomeres after the puncture of the embryo before transfer can be used for sex determination, preimplantative diagnostic, and marker-assisted selection in animal breeding.

## 8. Genetics in poultry reproduction

In birds and monotremes the sex is determined genetically but the exact mechanism is still unknown. All species show a ZZ/ZW sex chromosome system. Contrary to therian mammals it is characterized by female (ZW) heterogamety and male homogamety (ZZ). So far different hypothesis exist: Avian sex might be determined by dosage of a Z-coded gene or by a W-coded gene or may be both. A SRY gene or homolog is absent in birds and monotremes [22–24].

High performance chickens are bred by heterosis breed: In grandparent linages different traits (like high laying performance or good eggshell quality) are fixed by linage breed. In the parent generation the most suitable crossing between those linages will be determined. In case of the recurrent selection one father or mother line will be mated to a standard line. In the reciprocal recurrent selection each parent combination will be tested as mother and father line and rated according to the performance of the progeny.

Due to the chicken genome sequencing finished in 2005 many genes affecting vitality and disease resistance are going to be identified and become focused by the breeding organizations. This might lead to a reduction of therapeutic agents in poultry production.

## 9. Genetics in fish reproduction

Most of the fishes fostered in aquaculture belong to the teleosts. Many of them show hermaphroditism meaning that both sexes are present in one individuum but active at a different time during life. Most of the farmed fishes in aquaculture just release their eggs and sperms at the same time nearby in the water. In farms the eggs and sperms will be delivered artificially by abdominal massage and admixed together. This is the most suitable state for artificial manipulation like chromosome manipulation and gene transfer. Egg size is very important in fish breeding since large fry have to be fed later. Fertilization takes place by entering of the sperm through the micropyle of the fish egg, a narrow channel through the outer egg membrane. It is accomplished by the fish sperm swimming to the egg surface, locating the entry to the micropyle, and swimming down it to make contact with and penetrating the inner egg membranes. The membrane lifts after penetration of the first sperm which seems to be the mechanism for avoidance of polyspermy. Generation of triploid teleosts has become a widely used tool in order to protect wild fishes from hybridization since these animals are sterile. Production of monosex individuals has a similar effect [25]. Genetic engineering will have a great potential for the generation of disease resistance and low oxygen tolerating fish in the future. Interestingly somatic nuclear transplantation in teleosts leads to nucleocytoplasmic hybrids, meaning that the cytoplasm of the oocyte is able to inherit certain features. The molecular basis of this phenomenon is not elucidated so far [25].

## 10. Further perspectives of genetics in domestic animal reproduction

In many of the domestic animal species genome projects are ongoing or just finalized meaning that genes important for reproduction can be detected easier in order to investigate their specific role and the influences by the environment and metabolic state of the animal. In dairy cows classical selection has shown a negative correlation between daily milk yield and fertility parameters. In this regard, the metabolic state of the animal might be the limiting factor; therefore, the expression of reproductive genes under these conditions should be of high interest [26]. Modulation of gene expression is mediated not only by epigenetics (DNA methylation, histone modulation) but to a large extent also by micro-RNAs affecting mRNA stability. In this context the reproductive performance is influenced by the quality of the allele inherited, their epigenetic status, as well as the metabolic status of the animal.

In assisted reproduction sexing of sperms and embryos will become more important. Improvement of sexing technologies will allow a wider use also in domestic species which so far were excluded due to reproductive necessities (insemination doses, sensitivity of the sperm cells, survival time of the sperm cells). In embryos noninvasive techniques for sexing and preimplantative diagnostic may be possible in the near future. Selection of the sperm used for fertilization in intracytoplasmic sperm injection (ICSI) might change from motility alone to other, more complex criteria. *In vitro* culture conditions will be improved by adaptation to natural conditions avoiding aberrant gene expression leading to phenomenon like the

large offspring syndrome. In animal cloning the reprogramming of the donor cells will be improved in order to reach blastocyst rates around 20%.

In the social context the techniques applied in assisted reproduction should be rated in the correct light by understanding their way of action and their advantages in comparison to the objective risks.

## Author details

Sven Budik (Mag. rer. nat., Dr. rer. nat., Mag. med. vet.)

Address all correspondence to: sven.budik@vetmeduni.ac.at

Platform for Artificial Insemination and Embryo Transfer, University of Veterinary Medicine, Vienna, Austria

## References

- [1] Sex determination in mammalian germ cells. Spiller CM, Bowles J, Asian Journal of Andrology (2015) 17, 427–432; doi: 10.4103/1008-682X.150037
- [2] Molecular control of oogenesis. Sánchez F, Smitz J, Biochimica et Biophysica Acta – Molecular Basis of Disease (2012) 12, 1896–1912.
- [3] A tortoiseshell male cat: Chromosome analysis and histologic examination of the testis, Pedersen AS, Berg LC, Almstrup K, Thomsen PD, Cytogenetic and Genome Research (2014) 142, 107–111; doi:10.1159/000356466
- [4] The pseudoautosomal region and sex chromosome aneuploidies in domestic species. Raudsepp T, Das PJ, Avila F, Chowdhary BP, Sexual Development (2012) 6, 72–83; DOI:10.1159/000330627
- [5] A gene catalogue of the euchromatic male-specific region of the horse Y chromosome: comparison with human and other mammals. Paria N, Raudsepp T, Pearks Wilkerson AJ, O'Brien PC, Ferguson-Smith MA, Love CC, Arnold C, Rakestraw P, Murphy WJ, Chowdhary BP, PLoS One (2011) 6(7), e21374; doi: 10.1371/journal.pone.0021374. Epub 2011 Jul 25.
- [6] 46 XX karyotype during male fertility evaluation; case series and literature review. Majzoub A, Arafa M, Starks C, Elbardisi H, Al Said S, Sabanegh E. Asian J Androl [Epub ahead of print] [cited 2016 Dec 30]. Available from: <http://www.ajandrology.com/printarticle.asp?id=181224>
- [7] Xist localization and function: new insights from multiple levels. Cerase A, Pintacuda G, Tattermusch A, Avner P, Genome Biology (2015) 16, 166; doi: 10.1186/s13059-015-0733-y

- [8] Origin and evolution of X chromosome inactivation. Gribnau J, Grootegoed JA, *Current Opinion in Cell Biology* (2012) 24(3), 397–404; doi:10.1016/j.ceb.2012.02.004
- [9] Role of paternal and maternal genomes in mouse development. Barton SC, Surani MAH, Norris ML, *Nature* (1984) 311, 374–376; doi:10.1038/311374a0
- [10] Ligers and tigons and....what?....oh my! McKinnell Z, Wessel G, *Molecular Reproduction and Development* (2012) 79(8), Fm i; doi: 10.1002/mrd.22074.
- [11] Mammalian genomic imprinting. Bartolomei MS, Ferguson-Smith AC, *Cold Spring Harbor Perspectives in Biology* (2011) 3(7).
- [12] Genomic imprinting: parental influence on the genome. Reik W, Walter J, *Nature Reviews Genetics* (2001) 2(1), 21–32; doi:10.1038/35047554.
- [13] Significance of aquaporins and sodium potassium ATPase subunits for expansion of the early equine conceptus. Budik S, Walter I, Tschulenk W, Helmreich M, Deichsel K, Pittner F, Aurich C, *Reproduction* (2008) 135(4), 497–508; doi: 10.1530/REP-07-0298
- [14] Increasing expression of oxytocin and vasopressin receptors in the equine concepts between Days 10 and 16 of pregnancy. Budik S, Palm F, Walter I, Helmreich M, Aurich C, *Reproduction, Fertility, and Development* (2012) 24(5), 641–648; doi: 10.1071/RD11167
- [15] Pregnancy-acquired fetal progenitor cells. Seppanena E, Fiska NM, Khosrotehrania K, *Journal of Reproductive Immunology* (2013) 97(1), 27–35; doi: 10.1016/j.jri.2012.08.004
- [16] Vertebrate Reproduction. Blüm V, (1985) Springer, ISBN 978-3-642-71074-2
- [17] Equine fetal sex determination using circulating cell-free fetal DNA (ccffDNA). Moura de Leon PM, Campos VF, DellaCosta OA, Deschamps JC, Seixas FK, Collares T, *Theriogenology* (2012) 77, 694–698; doi: 10.1016/j.theriogenology.2011.09.005
- [18] The freemartin syndrome: an update. Padula AM, *Animal Reproduction Science* (2005) 87, 93–109.
- [19] Large offspring syndrome in cattle and sheep. Young LE, Sinclair KD, Wilmut I, *Reviews of Reproduction* (1998) 3, 155–163.
- [20] A New, Dynamic Era for Somatic Cell Nuclear Transfer? Loi P, Iuso D, Czernik M, Ogura A, *Trends in Biotechnology* (2016) pii, S0167-7799(16)30003-8; doi: 10.1016/j.tibtech.2016.03.008.
- [21] Mitochondrial DNA heteroplasmy in ovine fetuses and sheep cloned by somatic cell nuclear transfer. Burgstaller J, Schinogl P, Dinnyes A, Müller M, Steinborn R, *BMC Developmental Biology* (2007) 7, 141; doi: 10.1186/1471-213X-7-141
- [22] Avian sex determination: what, when and where? Smith CA, Roeszler KN, Hudson QJ, Sinclair AH, *Cytogenetic and Genome Research* (2007) 117, 165–173; doi:10.1159/000103177
- [23] Genetics of the Fowl: The Classic Guide to Poultry Breeding and Chicken Genetics. Hutt FB, ISBN: 978-0972177030

- [24] Poultry Genetics, Breeding and Biotechnology. Muir WM, Aggrey Stylus Pub Llc, (2003) SE, ISBN 0851996604
- [25] A Review of "Aquaculture and Fisheries Biotechnology, Genetic Approaches, Dunham RA, Stylus Cabi Publishing (2010) 2nd Edition," ISBN: 978-1-84-593651-8
- [26] Novel approaches to genetic analysis of fertility traits in New Zealand dairy cattle. Bowley FE, Green RE, Amer PR, Meier S, Journal of Dairy Science (2015) 98, 2005–2012; <http://dx.doi.org/10.3168/jds.2014-8266>



---

# **Studies on Gene Expression and Developmental Competence of Bovine Embryos Produced Under Different Conditions of Heat Stress**

---

Aya K. H. Mahdy, Krishna C. Pavani, Erica Baron and Fernando Moreira da Silva

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/66268>

---

## **Abstract**

Gene expression is required in all steps of embryonic development and therefore heat stress is known to reduce developmental competence after direct exposure of oocytes and embryos to different conditions of heat shock, by decreasing protein synthesis. Moreover, as in somatic cells, the heat stress befuddles the integration of RNA and posttranscriptional modification of RNA, the assumption was that during meiotic maturation heat shock may mutate RNA within oocytes, with the possibility of altering the surrounding cumulus cells, causing, thus, reductions in development. Heat shock proteins (HSP) are among the first proteins produced during embryonic development and are crucial to cell function. The HSP70 (*HSPA14* gene) is an important part of the cell's machinery for folding, unfolding, transport, localization of proteins and differentiation, regulation of the embryonic cell cycle and helping to protect cells from stress. Therefore, *HSPA14* is an apoptotic gene induced by heat shock is associated with embryonic loss, playing an important role of control mechanism of processes involved in growth, cellular differentiation, and embryonic development. In addition the connexin proteins (e.g. *Cx43*), related to gap junctions, are expressed in numerous tissues including gonads, act as a mediator of heat stress effect on cells. In the present review, the effect of heat stress on bovine embryonic development in a physiologic and genetic point of view is fully discussed.

**Keywords:** heat stress, oocyte, maturation, gene expression, *HSPA14*, *Cx43*, real-time PCR

---

## **1. Introduction**

Climate changes influence the biogeography and phenology of animals, thus affecting their reproduction, physiological development, and metabolism [1]. The global warming refers to

the continued increase of the earth's temperature and ecosystems change, as when the sun heats the earth, the earth radiates only some of the sun's energy into space, while some energy is trapped by atmospheric gases, such as carbon dioxide and water vapor. This energy builds up in the earth's atmosphere, the earth gradually becomes hotter and leads to increase in the temperature, and hence heat stress can happen [2]. Depending on the proportion of CO<sub>2</sub> emission, the global surface temperature is predicted to rise by 1.1–6.4°C. Heat stress is the main factor responsible for lower productive and reproductive performance in cattle during the summer months. The extreme temperature events due to climate change coupled with high rate of relative humidity decrease the ability of reproduction in animals [3]. Heat shock impacts on the ovulatory oocyte, as well as on functions of follicular granulosa and theca cells, therefore delays and alters the efficiency of follicle selection and lengthening of the follicular wave, which adversely affects the quality of oocytes [4]. The first-wave dominant follicle is depressed by heat stress; hence, lactating cows were found to be smaller in diameter. Thus during the first 7 days of the cycle showed that follicular dynamics have been altered, as indicated by the decrease in the number of medium-sized follicles that associated with depressed summer fertility [5, 6]. During hot summer months, high temperature can affect endocrine responses that may increase foetal abortions, shorten the gestation length, lower calf birth weight, and reduce follicle and oocyte maturation associated with the postpartum reproductive cycle. Postpartum heat stress can significantly decrease the pregnancy rates with impacts lingering well into the fall months [7]. The summer period adversely affects endocrine system, follicular phase, metabolism, function of oocytes, and embryos. The towering temperature causes several cellular changes during the maturation period and germinal vesicle stage in the oocytes [8]. Therefore, the heat stress influences and impacts on *in vitro* oocytes maturation, their nuclear maturation, and further embryos development to the stage of blastocysts after *in vitro* fertilization [9]. The low fertility of cattle females has led to an increased interest in *in vitro* embryo production (IVEP) technologies for achieving rapid genetic improvement and providing an excellent source of embryos for carrying out basic research on developmental physiology [10]. Moreover, cytoplasmic and molecular maturation of oocytes are thought to be critically involved in the ability of the oocyte to support fertilization and early development stages. The major activation of the bovine embryonic genome occurs at the 8- to 16-cell stage. Before embryonic genome activation, mRNAs, ribosomes, and proteins synthesized during oocyte growth and maturation contribute to early development [11]. As known, heat stress induces apoptosis or expression of *HSPA14* gene of *in vitro* produced embryos. Stress proteins are assorted into families depending on their molecular weight and provide two main functions: primarily as a molecular chaperone having key roles in folding/unfolding of proteins and secondarily as a stabilizer factor from deteriorating proteins contributing in the protection of cells against stress/apoptosis, granting an opportunity for rehabilitation or degradation in the cells suffering from cellular stress [12]. Some activities of HSP70 include folding, unfolding, transport, and localization of proteins and differentiation and regulation of the embryonic cell cycle [13]. Hence, the piling up of HSPs has long been considered a sign of cellular damage. There is an evident overlap of the signals that induces a protective stress response and those that initiate apoptosis [14]; therefore, heat shock protein HSP70 plays a protective role in the embryos. During the early 2-cell of embryo to the blastocyst stage, the HSP70 is the prevailing gene. The induction of stress proteins (HSP) synthesis during 1- or 2-

cell stage until blastocyst stage is intensely motivated by heat shock [15]. Thus, *HSPA14* gene is produced by embryonic cells, which protect the embryos from environmental stress. During early development, increased temperatures, free radicals, and oxygen stress have deleterious effects on embryonic viability and development. The *HSPA14* confers thermotolerance against variety of stressors. Thus, the preferential maintenance of *HSPA14* gene expression would allow this gene product to help maintain cellular function by acting as molecular chaperones to stabilize or refold proteins damaged by heat, and by blocking apoptosis [16]. There are many ways in which cells communicate with other cells, and gap junctional communication is one of these ways. Ovarian folliculogenesis and the production of fertilizable oocytes depend on gap junction channels which allow inorganic ions, second messengers, and small metabolites to pass from cell to cell, and this permeability is supposedly for underpinning the physiological roles played by gap junctions, for example, connexin 43 (*Cx43*) gene expressed within the oocyte–granulosa cell complex depending on the species [17]. The expression of *Cx43* gene in the ovary is regulated by gonadotropins like follicle-stimulating hormone (FSH) and luteinizing hormone (LH). Increase in the *Cx43* expression in the large antral follicles was observed where FSH was relatively elevated. Moreover, germline and somatic cell during mammalian ovarian follicle have elements for cellular interactions, development, and function. Somatic cell-to-oocyte communication is essential for oocyte growth and the regulation of meiotic maturation via gap junctions [18]. *Cx43* was later detected in the cytoplasm of germ cells. Therefore, tight junction molecules were found and *Cx43* were distributed at the cell–cell contact of the adjacent cells. Such progressive reorganization of germ cells, *Cx43* gene, is probably coupling with conjugation role of other gap junctions and their specific connexins (Cxs), between the different cell types that structure the animal reproductive system [19]. Furthermore, germline and somatic cell during mammalian ovarian follicle have elements for cellular interactions, development, and function, essential for oocyte growth and the regulation of meiotic maturation via gap junctions [23]. On the other hand, during embryonic early development, increased temperatures lead to free radicals and oxygen stress, having also deleterious effects on embryonic viability and development. As far as *in vitro* embryo production is concerned, more and more motives are actually presented in the cattle breeding, such as the faster propagation of superior germplasm because of the low efficiency of superovulation in programs of embryo transfer. Moreover, embryo production is actually also a reproductive method used to improve the number of progeny with high genomic merit from bovines, both *in vivo* and *in vitro*. *In vitro* embryo production (IVEP) also including some of the applications serves to solve infertility problems, genetic recovery, and production of clones or even transgenic animals. In the present review, physiological and genetic factors affecting embryonic development under heat stress conditions are fully discussed.

## **2. Ovarian physiology: folliculogenesis, oocytes growth, and development: oocytes growth and development**

The fertilization of mature oocyte occurs in the oviduct and there are three distinct phases that can be divided during this process. During the first phase (the oocyte growth phase),

the developmental competence of the oocyte and cell structure is generated, when oocyte growth accompanies follicular growth from the primordial to the small (2–3 mm) tertiary (antral) follicle. When follicles in a cohort reach a diameter of about 3–5 mm, one dominant follicle is selected during the antral phase. During the third phase, the oocyte undergoes to change (oocyte maturation) almost 24 hours between the peak ovulation and the rise of LH.

## 2.1. Oogenesis

The mature oocytes are differentiated, released, and established in mammalian ovaries to undergo oogenesis process for fertilization. The ovaries have individual follicles consisting of an innermost oocyte, surrounding granulosa cells, and outer layers of thecal cells, controlled by the endocrine system. Moreover, mammalian ovary produces steroids and peptide growth factors, which allow the development of female secondary sexual characteristics and support pregnancy [20]. The oocytes and surrounding granulosa cells have amiable connections together to support oocyte viability and growth mediated by the gap junctions, which are efficient conduits for low molecular weight substances. The granulosa cells have some metabolized molecules, which play a role in transporting the oocytes. Additionally, the KIT ligands and c-kit receptor are localized to oocytes and granulosa cells, respectively, to promote oocyte growth and follicular development. Moreover, some of the growth factors derived from oocytes, such as GDF-9 and BMP-15, contribute in follicular development by regulating the differentiation of surrounding somatic cells, as these communications are important for oocyte growth and follicular development [21]. The occurrence of events of oogenesis are concomitantly with folliculogenesis, as oogenesis can be explained as the process of formation and maturation of the egg by development and differentiation of the female gamete during the meiotic division, and this is the first phase of progress for the fetus. Thus, female ovaries have fixed number of oocytes, which decreases by time passing with several years, without potential to renew. During the ovarian start to deplete the auxiliary of oocytes, it evolves to a senescence/aging stage born with a fixed number of oocytes, which through several years reduce, without potential to renew. When the ovarian start to deplete the auxiliary of oocytes, it evolves to a senescence stage, leading the female to menopause [22]. During meiosis, mammalian oocytes undergo two consecutive asymmetric cell divisions, which are essential for the formation of a functional female gamete, without an intermediate replicative phase. Each division must ensure accurate segregation of the maternal genome and highly asymmetric partition of the cytoplasm, further that a tiny polar body and a large oocyte are generated. It leads to an asymmetric organization (or polarization) of the egg, which determines the geometry and the success of fertilization. Asymmetric divisions are tightly controlled by microtubule and microfilament cytoskeletons. During the beginning of the first meiotic division, this process allows the separation of the duplicated centrosomes and therefore to the gathering of a bipolar spindle formed by microtubules. Thus, it is referred to as the process that produces gametes with half of the number of chromosomes from the parent cells. For the position of spindle surrounding the oocytes and tossing the first polar body in parallel with separation of chromosome, the microfilaments are carried

out in meiosis I. The microtubule spindle is positioned at the surrounding of the oocytes until fertilization, causing the emission of the second polar body during second meiotic division (meiosis II). Additionally, the loss of asymmetry is a mark of low-quality oocytes and a signature of pre- and postovulatory aging [23].

## 2.2. Folliculogenesis

Meiotic prophase consists of several temporary stages: preleptotene, leptotene, zygotene, pachytene, and the diplotene stage in which the first meiotic division of the ovarian follicles begin to develop as primordial structure that oocyte arrested. Primordial follicle activation is characterized by the possession of complete layer of 11–20 granulosa cells around the oocytes. Secondary follicle stage starts to show features of a second layer of granulosa cells. Zona pellucida is the initial deposition material around the oocyte, and at the same time, cortical granules are formed within the oocyte cytoplasm. It is at this point of development that follicles appear to become responsive to gonadotrophins [24]. The progress of ovarian follicle is a combination of many sides of a compound process that starts with the foundation of limited pool of primordial follicles and attains in either atretic degradation of the follicle or liberation of mature oocytes for fertilization. Through the primary, preantral and antral stages, the primordial follicles must be grown during these stages before reaching to the pre-ovulatory stage where they are eligible to release oocytes for starting the fertilization stage. The corpus luteum structure (CL) is formed from the differentiation of the residual of granulosa and thecal cells after ovulation [25]. Large stock of oocytes is enclosed in primordial follicles in mammalian ovaries, and some of these follicles initiate growth toward a possible ovulation undergoing activity of the ovarian cycle. Additionally, most of these follicles end their growth at any moment and degenerate through atresia. During the growth of follicles, only a subset of oocytes is capable to support meiosis, fertilization, and early embryo development to the blastocyst stage. This proportion of eligible oocytes depends on the size of the follicular cell. Developing lines of evidence propose that the competent oocytes increase the storage of gene production leading to the determinant to support the precocious stages of developmental embryos, before the activation of embryonic genome. Thus, these transcripts may be stored during early folliculogenesis as the oocyte grows and displays high transcription activity [26]. Young and McNeilly [27] classified the system into five types: type I represents the primordial follicles, which are a resting stage before their activity begins. In this stage, follicles have only one layer of granulosa cells and this is the follicle transitioning through the primary stage, when the granulosa cells become cuboidal. The second type of follicles includes one layer of cuboidal granulosa cells. While antral follicles include two to four layers of granulosa cells going to the third type. At this stage, the large preantral follicles consists of four to six layers of granulosa cells, considering this stage as the fourth stage, increasing the number of layers, thus reaching the fifth type. Then, the *antrum* arises and the thecal cells start to appear and protrude. Thus, begins the formation of the layer around the granulosa cells of the oocyte. Therefore, most follicles are observed at early stages of development. At the antral stage, follicles become gonadotropin dependent and form large antral follicles, most of which undergo atresia, and few are selected for ovulation. Estrogens

produced from the ovary of vertebrates, which have a fundamental endocrine function, leads the females to develop reproductive organs, while the corpus luteum produces progesterone, which is essential for the foundation of pregnancy. Thus, these functions are tightly coordinated during folliculogenesis, in which a dynamic process includes a continuous differentiation of three types of cells, theca of granulosa cells, and the oocytes themselves. The antrum formation from the granulosa cells proliferate, which is directly related to the follicular growth leading to a restricted number of follicles to complete their development to the stage of ovulation, while others undergo atresia. In the stage of antral follicle, theca cells are divided into two layers, internal and external, which can be distinguished by the surrounding selected follicles [28]. After ovulation, which is triggered by a peak of LH, theca cells and mural granulosa cells luteinize to produce progesterone (Figure 1). Once an ovarian follicle increase from the pool of resting follicles and initiates growth, it takes approximately around 100 days to arrive to the point where ovulation can execute and the oocyte included within the follicle is released [29].

### 2.3. Oocytes morphology and classification

*In vitro* maturation (IVM) is the starting point of a whole lot of biotechnological applications in animals like *in vitro* fertilization (IVF). These techniques combined with marker-assisted selection at embryonic stage will hasten acceleration for improving the production potency of cattle. The selection of oocytes depends only on the good quality cumulus-oocytes complexes (COCs), which were selected based on cumulus cells and ooplasm characteristics. The good oocytes are surrounded with at least three layers of cumulus cells and their ooplasm should be dark and homogeneous. Therefore, the oocytes were classified into four classes: (1) Grade A: oocytes were characterized with more than five

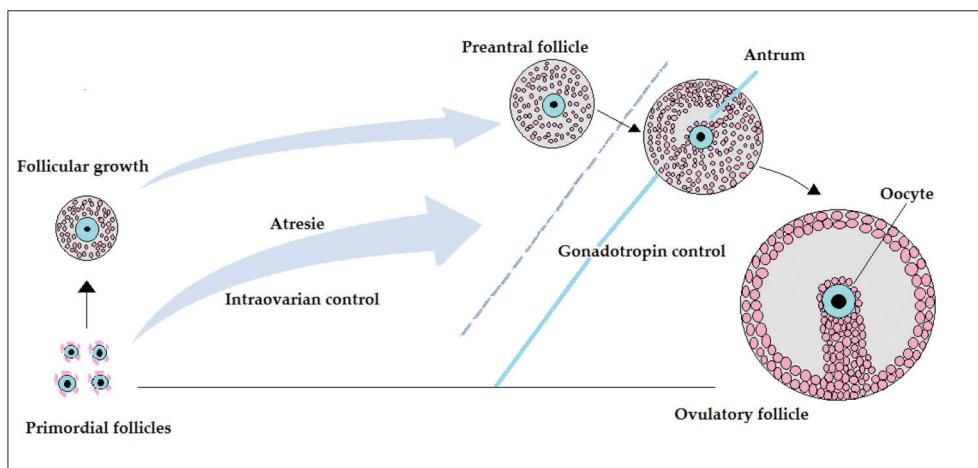


Figure 1. Outline of the main steps of folliculogenesis.

complete layers of cumulus cells, which are uniform granulation of ooplasm and healthy follicle. (2) Grade B: follicles with intact and well-organized granulose cells with three to five complete layers of cumulus cells and uniform granulation of ooplasm. (3) Grade C: follicles with one to two complete layers of cumulus cells and uniform granulation and less regular of ooplasm with some dark area. (4) Grade D: denuded oocytes with uniform granulation of ooplasm [30].

## 2.4. Oocyte maturation

Immature oocytes begin to develop in the ovaries, with each oocyte possessing a large nucleus that is referred to as the germinal vesicle (GV). Therefore, immature oocytes start to give response and undergo maturation process, which starts the nucleus of oocyte to have disassembly during a sequence called germinal vesicle breakdown (GVBD). After this stage, the immature oocyte turns into mature being eligible for developing until fertilization. Moreover, regardless of the timing of GVBD relative to fertilization and at the end, all oocytes must be matured enough to be able to consequently develop to continue to proceed normally [31]. During dictyate stage of prophase I, the oocytes are arrested, which can be identified by the presence of a germinal vesicle (GV). Additionally, meiosis I is marked by germinal vesicle breakdown (GVBD) following which bivalents are brought to alignment at the spindle equator by metaphase I. Anaphase I then ensues when chromosomes segregate between the secondary oocyte and the polar body. Following first polar body extrusion (PBE), oocytes progress without a hiatus into meiosis II where they are arrested for a second time at metaphase II [32]. During the maturation process, mammalian oocytes accompany a comprehensive, extensive rearrangement of the cytoskeleton and associated proteins. In oocytes, during the MI takes nearby 6-11 hours and the spindle migration toward the egg cortex occurs at this time. When the chromosome-spindle complex moves to the egg cortex, it involves a spindle pole close to the cortex [33]. Spindle movement induces a cortical differentiation performed by the accumulation of actin filaments and a scarcity of microvilli. After polar body extrusion, chromosomes realign progressing to metaphase II.

### 2.4.1. Nuclear and cytoplasmic maturation

The oocytes for bovine with an inside zona diameter smaller than 95  $\mu\text{m}$  are unable to start meiosis *in vitro*. A high percentage of bovine oocytes are able to start meiosis to the MI stage once the oocyte diameter is at least 100  $\mu\text{m}$ . However, the oocyte must measure 110  $\mu\text{m}$  or more to reach the MII stage. The ability to develop to the blastocyst stage *in vitro* increases with oocyte growth. Cleavage and blastocyst rates increase in parallel with meiotic competence and significantly higher developmental rates have been obtained when the diameter of fertilized oocytes is greater than 120  $\mu\text{m}$ . The developmental potential is apparently similar in oocytes originating from nonatretic and early atretic follicles [34], including spindle transfer, chromosome condensation, germinal vesicle breakdown, the progression to MI, and separation of the homologous chromosomes with polar body extrusion, where the nuclear

membrane starts to fold, the nuclear pores disappear, and then the nuclear membrane undergoes fragmentation and rapidly disappears. It appears that nuclear maturation follows the same pattern *in vivo* and *in vitro*. Nuclear maturation involves changes in protein synthesis patterns. Bovine oocytes undergo marked changes in the patterns of protein synthesis after GVBD *in vitro* and *in vivo*, whereas oocytes that remain at GV stage have consistent protein synthesis patterns [35]. Meiotic competence is known as the ability of the oocyte to complete meiosis. Additionally, during follicular growth, the meiotic competence is obtained and is acquired progressively. Oocytes first acquire the capacity to undergo condensation of chromosome, germinal vesicle breakdown, and follicular growth and are desired to acquire the capacity to progress to MI and eventually acquire the capacity to reach MII. The ability to complete the MI to MII transition coincides with the achievement of full size and with the process of nucleolar compaction [36].

The diameter of the follicle increases depending on the number of Golgi apparatus present in the oocyte. The change in location of cortical granules constitutes the most obvious ultra-structure sign of cytoplasmic maturation. According to the oocytes in the GV stage, cortical granules are distributed in clusters throughout the cytoplasm. However, as the oocytes progress to metaphase I stage, the cortical granules translocate to the periphery of the oocyte and become attached to the plasma membrane. At the end of the maturation period, when these oocytes reach the MII stage, the granules are distributed through the inner surface close to the plasma membrane [37]. The mitochondria make a homogeneous distribution throughout the cytoplasmic and are more common at the germinal vesicle (GV) stage, while heterogeneous distribution is more commonly observed in the oocyte of metaphase I or II. During oocyte maturation, the mitochondria disperse distribution throughout the cytoplasm, until reaching metaphase II (MII), when the central position in the cell operates in the mitochondria because high-energy supply around the nucleus is very important during embryonic development. It is also observed that morphologically poor quality embryos are only characterized by the homogeneous distribution of the mitochondria [38]. Additionally, the cytoplasmic maturation describes both the ultrastructural changes that take place in the oocyte from the germinal vesicle (GV) to the metaphase II (MII) stage and the possession of developmental competence of the oocyte. Mammalian oocyte's cytoplasmic maturation can be described as the ability of a mature egg to undergo regular fertilization stage, all stages of cleavage, and further development of the blastocyst. Other indirect morphological parameters considered to evaluate cytoplasmic maturation include cumulus cell expansion, polar body expulsion, and increased perivitelline space (PVS) of Ref. [39].

### 3. RNA synthesis and molecular maturation

The molecular maturation coincides with the maturation and growth of oocyte corresponding with transcription and mRNAs expression by genes of oocytes. In mammals ovaries, meiosis occurs mainly in two steps: during foetal life (step 1) and through the period of preovulatory life (step 2). The genetic information must be stored in advance because of nuclear instructions, when the chromosomes are the subjects of transformation. Thus, in the

form of mRNA, the chromosomes arrive and are silenced at the resumption of meiosis. After DNA synthesis at double helix, the chromosomes are partially condensed and rearranged by the process of crossing over. The chromatin then reaches a special conformation that is an intermediate between chromatin condensation and interphase, the dictyate phase. Additionally, in some species during that special prophase period, the oocyte remains static and the chromosome appearance changes little. When an oocyte begins to grow in the primordial follicle, the uncondensed loops of chromatin in the dictyate state ensure the transcription of required elements. Thus, the mRNA produced is either translated immediately [40]. Through the *in vivo* and *in vitro* maturation, oocytes can have the ability to elucidate the signaling pathways by gene expression. These signaling pathways are involved in the intricate mutual among the oocyte and its somatic compartment through morphogenetic and differentiation processes, and the origin of disturbances in oocyte maturation is theoretically involved in the decrease of the fertility; therefore, molecular maturation of oocytes is characterized by enhanced nucleolar activity and ribosome synthesis, and when the oocyte reaches the meiotic phase, the nucleolus becomes inactivated [41]. During the meiotic phase, for oocyte maturation, large numbers of macromolecules are accumulated. The mRNA and rRNAs produced in these cells are far in excess of those necessary to support protein synthesis. Thus, concomitantly the onset of germinal vesicle breakdown occurs after 6–10 h of IVM. In the course of IVM, *de novo* transcription strongly declines as determined by measuring the incorporation of [<sup>3</sup>H] uridine into RNA. In contrast to this finding, the incorporation of [<sup>3</sup>H] adenosine increased and showed a peak during the time interval of 6–10 h of IVM, parallel with the onset of germinal vesicle breakdown (GVBD) and translation. In the further course of maturation, only a moderate decrease of [<sup>3</sup>H] adenosine incorporation was observed. These findings indicate and betoken that during the time of germinal vesicle breakdown (GVBD) the translation is increased. Additionally, these operations were accompanied by polyadenylation of the mRNA, despite the decline of the transcription and accumulation of polyadenylated mRNA until MII [42]. As known, oocytes' growth normally undergoes cytoplasmic and nuclear maturation, and the bovine embryo has the ability to develop to the blastocyst stage, and it is important to clarify the contribution of the oocyte to the embryo quality. Currently, the most popular hypothesis is that specific mRNA and possibly some proteins are produced and added to the oocyte's stockpile in the last few days before ovulation, it is believed that molecular maturation represents the closest association with the intrinsic capacity of an oocyte to reach the blastocyst stage and probably beyond [43]. Gene expression patterns are responsible for the development of the early embryogenesis. Fertilization and the first zygotic cleavage involve major changes to paternal and maternal genome activation. Additionally, the function of genomes has regulatory mechanisms, including differential promoter activation, alternative RNA splicing, RNA modification, RNA editing, localization, translation and stability of RNA, expression of noncoding RNA, antisense RNA, and microRNAs. All of these mechanisms work together to produce the level of RNA for transcriptome functioning of an organism. Developmental embryonic stage is the first to rely on the stored maternal transcript, which is progressively exhausted until the production of the embryo on its own transcripts, after shifting to the program for embryonic expression. Moreover, through gene expression patterns and RNA stability, oocytes and early embryos prior to the expression of embryonic activation are dramatically

different and it leads to significantly different from what is observed after the main outset of embryonic transcription. The start for transcription of embryonic genes execute at a specific species at a specific time point. In bovine embryos, it occurs at 8- to 16-cell stage [44]. When the immature oocytes are removed from follicle, they spontaneously will convert to mature oocytes. Oocyte maturation can have deficiencies and the incompetence of cellular machinery, frequently causing a failure in embryonic development following fertilization and lower implantation rates. Additionally, the cytoplasmic and molecular maturation of oocyte might fail to promote male pronuclear formation. Hence, it might increase the chromosomal abnormalities after fertilization, and these abnormalities may result in an incompetent embryo in cleavage [45].

#### 4. Production of embryos *in vitro*

There are many reasons for interest in *in vitro* generation embryos that can be produced in the laboratory. Among these reasons is the faster propagation of superior germplasm in cattle because of the low efficiency of superovulation (SO) and embryo transfer (ET) programs. *In vitro* embryo production (IVEP) consists of four steps: (1) aspiration of the ovaries, (2) *in vitro* maturation (IVM) of the recovered oocytes, (3) *in vitro* fertilization (IVF) of the IVM oocytes, and (4) *in vitro* culture of the IVM/IVF zygotes for development to the desired stages [46]. To evaluate the effect of oocyte source (live animals and abattoir ovaries) on subsequent embryo development in cattle, it is essential to perform IVEP from cow ovaries. The beginning of harvest of the oocytes from bovine ovaries and the cumulus-enclosed oocytes (COCs) suitable for IVEP were *in vitro* matured (IVM), fertilized (IVF), and cultured (IVC) to the tight morula (Tm) and blastocyst (Bl) stage. A higher overall IVEP efficiency is mainly related to the higher cleavage rate. Moreover, the production of embryos *in vitro* is a technique largely used as a method for increasing the production of progeny in the bovine with high genetic meritocracy, overcoming several of infertility problems in cattle, genetic recovery, and production of clones and transgenic animals. Despite several studies in this area, the *in vitro* development of cattle embryos to the blastocyst stage rarely surpasses the rate of 40%. Extrinsic and intrinsic oocyte factors can interfere with oocyte fertilization and development to the blastocyst stage. Among the intrinsic factors is the presence of the dominant follicle that, mainly through inhibin and estradiol secretion, has an inhibitory effect on the development of other follicles. Furthermore, they reported that oocytes derived from subordinate follicles have less capacity to become healthy embryos than oocytes derived from growing follicles [47].

In **Figure 2**, it can be observed different stages of embryonic development from day 0 to day 7, monitored by our team. Following fertilization, embryos undergo a series of mitotic cell divisions. Hence, the embryo compacts to form a morula that comprises of cells in a compact cluster including the pellucid zone (i.e., comprising of glycoproteins envelope of mammalian oocytes, which beset the embryo). Then, the blastocyst is formed and finally “hatches” from the zona pellucida. Besides, in humans, all this process takes about 1 week [48], and in cows, it can take up to 10 days.

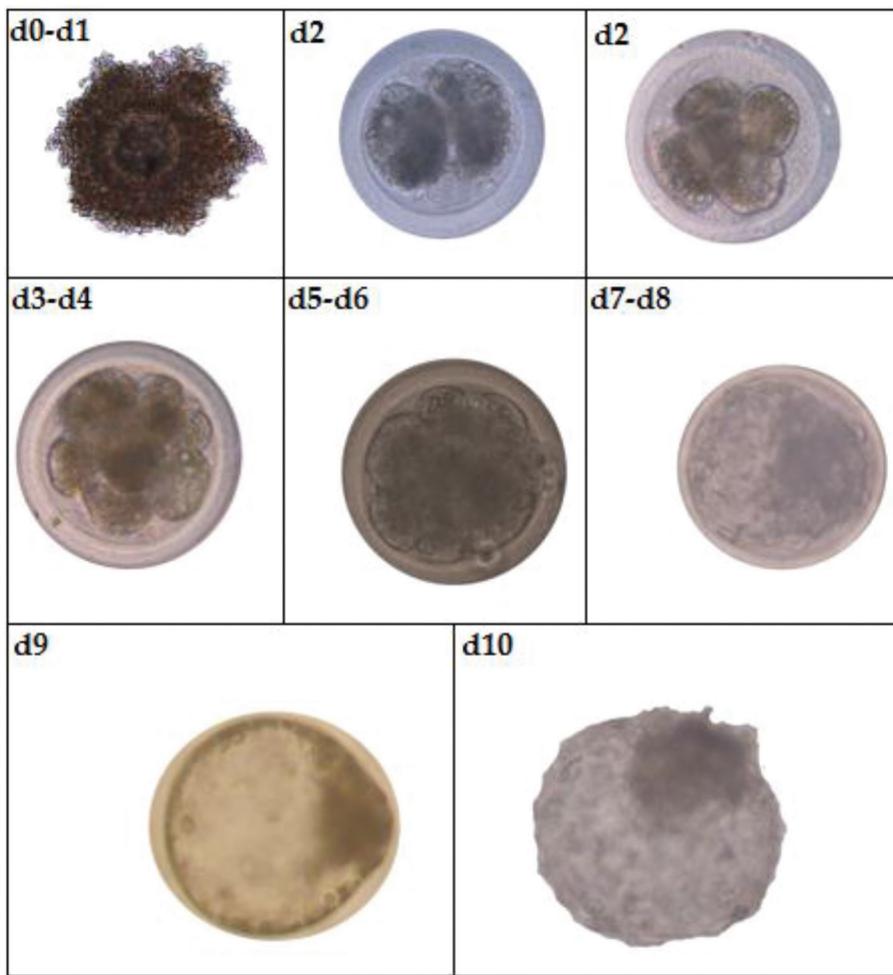


Figure 2. Stages of embryo development (photos made by the authors).

## 5. Effects of heat stress on animal fertility

During summer, heat shock reduces pregnancy, conception rates, and leads to low fertility in lactating dairy cow. Progesterone secretion by luteal cells is decreased and this is reflected in plasma progesterone concentration, the endometrial function, and alters its secretory activity by high temperature, which leads to the termination of pregnancy. Moreover, heat shock impedes oocytes quality and embryo development and increases the mortality rate of embryos [49]. Therefore, there are different factors that lead to the decrease of fertility in cattle. The most important are the increased humidity and temperature that result in a decreased expression of overt estrus and a reduction in appetite and dry matter intake.

During the postnatal period “puerperium,” the exposure to extreme heat stress prolonged, leading to negative energy balance and increase in the calving pregnancy interval. Heat stress influences by decreasing the dominance of the selected follicle and this decreases the steroidogenic ability of its granulose cells and theca cells, which declines the estradiol concentrations in blood. The elevation or reduction of the plasma progesterone levels can be relied on whether it is severe or chronic, and on the cattle metabolism state. The reduction of follicular activity and the alteration of ovulatory pattern execute endocrine changes, leading to the reduction of oocytes and the quality of embryos [50]. Moreover, oocyte's maturation from the meiotic stage of germinal vesicle (GV) to MII is affected by heat stress and it has been shown that high temperatures affect also the cattle cumulus-oocyte complexes (COCs), zona pellucida hardening, fertilization, and further cleavage of putative zygotes after fertilization [51]. In cattle, genes exist for regulation of body temperature and for cellular resistance to elevated temperature. These genes offer possibility for their incorporation into dairy cow through cross breeding or on an individual-gene basis and the physiological and genetic manipulation of the cow to improve embryonic resistance to high environmental temperature, which reduces fertility in lactating dairy cows, and as a result, pregnancy, oocyte, and early embryo are affected by heat stress [52]. Therefore, heat stress reduces gonadotropin secretion and the ovarian pool of oocytes and impairs fertility. Additionally, heat stress has immediate effects on follicle and its enclosed oocyte, follicular function, and follicular growth and disrupts steroidogenesis [53]. Heat stress has an impact on the reproductive function, that is, it reduces the intensity of the behavioral estrus and leads to low fertility in female and compromised sperm output and increased sperm abnormalities in male. The pregnancy rate and the embryos at earlier stages of development during heat stress are affected [54]. Furthermore, oocyte susceptibility to heat shock can be detected during the germinal vesicle (GV) and oocyte maturation periods. The bovine oocytes' exposure to high temperature *in vivo* and *in vitro* affects oocyte maturation, fertilization, and preimplantation of embryonic development. This heat-induced reduction in oocyte function occurs due to a series of cellular alterations that affect nuclear and cytoplasmic compartments of the bovine oocyte. Before ovulation, there are several physiological factors that can disrupt establishment of pregnancy changes by elevated thermal heat. Heat shock effects endometrial prostaglandin secretion, oocytes during maturation period, and decreases low fertility of dairy cows in hot seasons [9]. Moreover, heat stress was suggested to be similar to oxidative stress, because of correspondences in the genes expressed after heat exposure, such as genes encoding heat shock proteins (*HSPA14* gene), demonstrated heat-induced increase of reactive oxygen species (ROS) production, especially the superoxide anion. Additionally, heat stress disturbs protein synthesis, increasing the production of heat shock proteins (HSPs). Due to their chaperone function, HSPs ensure the folding, unfolding, and refolding of nascent or stress-denatured proteins. The HSP70 and HSP90 were correlated with the development of thermotolerance of embryos. Thus, heat stress was shown to increase HSP70 and HSP90 levels and that probably due as a mechanism of cellular defense and/or repair [55]. Dairy cows respond to heat stress in several ways, including reduced feed intake, lower milk yield and quality, and compromised fertility [56]. Thus, fertility traits in dairy animals show a very low heritability value, and this indicates that most of the variations in the fertility are determined by nongenetic factors or environmental effects. Therefore,

summer season has been associated with reduced fertility in dairy cattle through its deleterious impact on oocyte maturation and early embryo development. Moreover, the highest pregnancy rate of cow was observed in September–November as 32% while the lowest pregnancy rate of 24% in March–May. The lower pregnancy rate is due to the delay of rebreeding cows in the summer hot months with a high level of heat stress. Additionally, there are some other genes such *ATP1A1* allowing resistance and adaptation to thermal stress in bovine. Whereas *ATP1A1* gene is known as  $\text{Na}^+/\text{K}^+$  -ATPase subunit alpha-1, this gene is fully recognized as a nominee to respond for heat shock because of its assembly to oxidative stress in bovine which have  $\text{Na}^+/\text{K}^+$  -ATPase protein complex consisting of  $\alpha$  and  $\beta$  subunits. The *ATP1A1* gene encodes the  $\alpha 1$  isoform, for  $\alpha$  subunit of  $\text{Na}^+/\text{K}^+$  ATPase pump is considered as a major isoform. *ATP1A1* gene has been mapped on *Bos taurus* chromosome number 3 and is comprised of 22 introns and 23 exons. *ATP1A1* gene is responsible for the fundamental establishment of the electrochemical gradient of  $\text{Na}^+$  and  $\text{K}^+$  across the plasma membrane, which is fundamental for maintaining body fluid and cellular homeostasis. During the elevation of body temperature “hyperthermia,” heat stress activates heat shock transcription factor-1 and promotes the expression of HSPs coupled with reduced expression and synthesis of other proteins having a fundamental role to activate the immune and endocrine system. For understanding of genes in the regulation of heat shock response in animals would be helpful to improve their thermal tolerance via gene manipulation. Therefore, the *HSPA14* genes were found highly expressed in summer months in cattle, which enhance their thermotolerance and the ability to adapt to the thermal environment [57]. Genetic adaptation to heat stress for cellular resistance to elevated temperature reduced the levels of gonadotropin receptors, aromatase activity of granulosa cells, the follicular fluid concentrations of oestradiol, and follicular function involving changes at the level of the follicle or the secretion of the pituitary hormones that control development of the follicle. Additionally, oxidative stress is a major cause for thermal damage of spermatogenic cells and leads to apoptosis and DNA strand breaks. Moreover, apoptosis plays a critical role in effects of thermal stress on the maturing of oocyte in cattle. Therefore, the inhibition of apoptosis in bovine embryos with a caspase inhibitor increased the magnitude of the reduction in development caused by the elevated temperature. Thus, apoptosis is limited to the most damaged cells of the embryo. Heat shock caused damage to the oocyte during preovulation, which seems to involve the generation of reactive oxygen species, as both are effects of heat stress *in vivo* and heat shock *in vitro*. Furthermore, there are indications that developmental competence of the resulting embryo can be reduced if fertilization is by a spermatozoon exposed to heat shock. *In vitro* fertilization with sperm recovered from male in which the scrotum was heated to 42°C resulted in embryos with reduced ability to complete development [58]. In cows, heat shock has negative effects on competence of oocytes from antral follicles, granulosa, or theca cells, which are responsible for the production of steroids. Moreover, heat stress applied on bovine oocytes during *in vitro* maturation as well as on embryos during *in vitro* culture reduces both oocyte maturation and embryo development rates. The expression of HSP70 was not affected in the oocyte, but increased in cumulus cells. Therefore, the negative effects in oocytes because of heat shock could be mediated through the surrounding cumulus cells. Interestingly, higher expression of HSP70 in immature bovine oocytes was observed. This mRNA could be translated throughout the IVM and consequently decreases its expression

in mature oocytes [59]. Heat stress accelerates cellular metabolism, resulting in the inability of the mitochondria to properly reduce oxygen that remains in its radical state [5, 6]. Thus, the poor developmental competence might be due to the lack of cytoplasmic maturation in oocytes maturing *in vitro*, even though they undergo normal nuclear maturation. In addition to the presence of heat shock proteins, there are several antioxidants, for example glutathione (GSH) that partially imparts the protection against oxidative stress. GSH appears to be the main antioxidant defense system against reactive oxygen species (ROS) in oocytes/embryos [60].

## 6. Regulation of the cell cycle and oocyte maturation

Several mechanisms are involved in the activation of translationally inactive mRNA. These mechanisms involve the phosphorylation of many factors that initiate translation. Therefore, according to this model, polyadenylation (the addition of adenine) of the 3' terminal portion of the cytoplasmic mRNA would stimulate the release of repressor molecules linked to the 5' portion, thus beginning translation. The transport of this mRNA to the cytoplasm occurs through a characteristic splicing of the poly-(A) tail, which, after reaching the cytoplasmic compartment, becomes smaller and heterogeneous in size. The cytoplasmic elongation of the poly-(A) tail has been associated to the translation activation, meaning that during addition of adenine to mRNA in the cytoplasmic of oocyte through maturation leads to deadenylation, leading to the degradation of the particular mRNA. For this reason, protein synthesis starts when the two ribosomal subunits are linked onto the mRNA. This stimulates the degradation of particular mRNA which, when oocytes are acquiring developmental competence, the fundamental transcripts produced encode regulators of the cell cycle such as "maturation promoting factor" (MPF), the protein of the c-mos pro-oncogene (MOS), and mitogen-activated protein kinase (MAPK) [37]. In mitotic cells, S-phase always precedes M-phase in order to maintain euploidy. Additionally, the maturation (M-phase) promoting factor (MPF) plays a pivotal role in oocytes during their maturation. During the mitotic cell cycle, MPF activity shows different stages of vacillation and steadiness, i.e., MPF activity wobbles positively and negatively in time with the beginning and ending of M-phase in succession, while being precisely regulated during the two cell divisions of meiosis. MPF is a heterodimer protein kinase complex consisting of two subunits: the catalytic subunit CDK1 and the regulatory subunit cyclin B1. As phosphorylation is required by CDK1, cyclin B is the main determinant for CDK1 and this is a necessary factor for certain processes during cell cycle as (1) initiating germinal vesicle breakdown (GVBD) through phosphorylation of nucleoporin—a components of nuclear pore complex—among many protein components of nuclear envelope and (2) with a large protein complex termed as condensing that helps in supercoiling the DNA during mitosis. The kinase activity requires more than binding of CDK1 to cyclin B1. MPF, the heterodimer protein, is held in an inactive state termed as PRE-MPF by wee1 kinase that causes inhibitory phosphorylation of CDK1 (p34<sup>cdk2</sup>). However, CDC25 activates CDK1 by dephosphorylation at the same sites. Although binding

of CDK1 to cyclin B1 is necessary, it is not sufficient for kinase activity. Switching on MPF in all cells is further governed by the balance in the regulatory activity of Wee1/Myt1 kinases, which cause an inhibitory phosphorylation of CDK1 (p34<sup>cdk2</sup>) (and hold the heterodimer in an inactive state called pre-MPF), and the CDC25 phosphatases, which cause activation of dephosphorylation of CDK1 at the same sites. Thus, the high CDC25 and low Wee1/Myt1 activity are needed for switching on the CDK1 component of MPF. In addition, before entry into mitosis, cyclin B1 (and so MPF) is spatially restricted to the cytoplasm, through a cytoplasmic retention sequence, containing a nuclear export signal. When cells commit to mitosis, cyclin B1 has to become phosphorylated within its cytoplasmic retention sequence, leading to rapid accumulation of cyclin B1 and MPF within the nucleus, and ensuing GVBD [61]. In the mammalian oocytes during the meiotic maturation, two consecutive divisions execute without an intermediate phase of DNA replication where even haploid gametes are produced. In addition, these two divisions are asymmetrical, maintaining and ensuring the maximum number of oocytes possessed for maternal stores. In addition, immature oocyte contains only a small amount of cyclin B, just enough to induce entry into the first meiotic M-phase. Furthermore, MPF activity is regulated by a translation-dependent mechanism that determines the level of cyclin B1 synthesis. Among the different mechanisms that control the expression of maternal mRNAs, polyadenylation has been implicated in cyclin B1 translation in *Xenopus* and mouse oocytes. The immature oocyte contains only a small amount of cyclin B1, just enough to induce entry into the first meiotic M-phase increases progressively, reaching its maximum at the end of the first meiotic M phase, and the newly synthesized protein becomes associated immediately with the p34<sup>cdk1</sup> kinase to form an active complex. Cyclin B degradation is required for polar body extrusion. Changes in cyclin B1 levels, through changes in MPF activity, regulate not only the timing of the cell-cycle phases during meiosis but also the orderly events leading to the formation of functional meiotic spindles and asymmetric divisions, MPF activity controls the formation of a functional spindle in the oocyte. The MPF activity required for GVBD (sufficient for entry into M-phase) only allows the formation of a single aster of microtubules around the condensed chromosomes. A first threshold in MPF activity is then required to organize the microtubules into a bipolar structure. In contrast, the further migration of the chromosomes toward the vicinity of the spindle equator does not depend on changes in the MPF level [62]. The meiotic maturation is an essential process for the development of an immature oocyte into fertilization egg. In vertebrates, it is compatible with the transition from the prophase arrest of the first meiotic division to the metaphase arrest of the second meiotic division. Thus, MPF is activated in response to the hormonal signal. Therefore, steroids bring about meiotic maturation through functionally redundant pathways involving synthesis of Mos or of cyclin proteins. The c-mos pro-oncogene (MOS) protein is a Serine/Threonine kinase, which is specifically expressed in germ cells where it functions only during the short period of meiotic maturation from maternal mRNA in vertebrates before being proteolyzed at fertilization. Although it is not necessary for GVBD, Mos remains a powerful inducer of meiotic maturation when microinjected. Two roles could be envisioned for Mos in the process of GVBD induction: (1) The mRNA coding for Mos could, intrinsically or via associated proteins, affect the translation of other mRNA. (2) Synthesized Mos kinase provokes a

strong pathway that most certainly participates in the amplification of MPF, as well as being necessary for normal GVBD events (for the repression of DNA synthesis) [63]. Mammalian oocytes reach the prophase diplotene stage of meiosis I before or after birth, and they remain arrested at this stage until resumption of meiosis, characterized by germinal vesicle breakdown (GVBD) that follows preovulatory gonadotropin stimulation after puberty. Thus, family of serine and threonine kinases is activated in the somatic compartment of the follicle. Intra oocyte MAPK cascade activation is more closely related to post-GVBD events such as meiotic spindle organization. Stimulation of meiosis resumption by activation of MAPK can be accomplished by prompting synthesis of downstream meiosis resumption inducing factors. Additionally, the phosphorylation of gap junctional proteins blocks communication between oocyte and around somatic cells, and subsequently prevents the meiosis from inhibiting signals from entering into oocytes from the vicinity of somatic cells. Lutropin causes activation MAPK in follicular somatic cell, which have fundamental role, which in turn phosphorylates connexins, leading to a reduction in gap junction permeability between the somatic cells prior to germinal vesicle breakdown. The suppression of gap junction communication is enough and essential for reinitiating meiosis. Mos kinase is a universal mediator of oocyte meiotic maturation and is produced during oogenesis and destroyed after fertilization. Thus, mos limits the number of meiotic division to just two rounds (meiosis I and II) and to avoid from ingress into meiosis III. Strikingly, the maintaining of Mos/MAPK pathway activities after fertilization close to physiological levels prompts additional rounds of meiosis, and the spindle is positioned symmetrically resulting in further rounds of asymmetric cell division [64]. Additionally, in the mammalian oocytes the meiotic maturation is described by special asymmetric cytokinesis through chromosomes moving from the center toward the cortex zone of an oocyte. After GVBD, a meiotic spindle congregates surrounding centrally positioned metaphase chromosomes and then moves in the vicinity of the cortex in an actin filament relying on the process. Furthermore, cortical reorganization executes an ectopic actin-rich actin cap and a cortical granule-free domain (CGFD) formed in the vicinity of the cortex devoid of microvilli in a MOS-dependent manner. Establishing cortical polarity, including spindle migration, positioning and cortical reorganization is critical for oocyte asymmetric divisions. Then cytokinesis occurs and the polar body extrudes, forming a high-polarized MII oocyte [65]. Additionally, in humans, oocyte growth unfolds over 110–120 days, during which cell mass undergoes an astounding, more than 100-fold, increase and cell diameter shifts from <40 to ~120 mm. At the same time, the macromolecules and organelles are produced and stored in very large amounts. In such a way, the oocyte meets a fundamental for preimplantation embryo developmental growth, storing enough for cytoplasmic mass desired for fertilized egg to accomplish multicellularity without the presence of net growth. Remarkably, the oocyte growth is much more than a quantitative increase in cell mass finalized to the housekeeping needs of the early embryo. For example, maternal effect genes have been characterized as sequences whose transcriptional and translational products are generated during oocyte growth and whose regulatory function is required only after fertilization. In particular, many maternal effect gene products are transcriptional regulators whose loss of function causes major developmental failures, such as arrest at cleavage or blastocyst stages, or inability to activate the zygote genome [66].

## 7. Regulation of mammalian oocyte gene expression at transcription level

There are six Obox (oocyte-specific homeobox) family transcripts and Obox-1, 2, 3, and 5 mRNA they have been detected in oocytes from growing primary follicles [67], playing then an important role in early embryogenesis [68, 69] by orchestrating gene transcription, either ubiquitously or in a tissue-specific manner. Mice lacking the Obox6 gene grow without morphological abnormalities and with normal fertility, indicating a functional redundancy among the Obox family members [70]. Moreover, several genes play key roles in oogenesis, folliculogenesis, or early embryonic development. In particular, GDF-9 and BMP-15 are necessary for folliculogenesis beyond primary follicles in mouse and sheep, respectively. Gene expression in oocyte is quite different from those in somatic cells. The messenger RNAs produced by these cells are not only required to support germ cell development but, in the case of oocytes, they are also used for maturation, fertilization, and early embryogenesis. It is very important to understand the oocyte mechanisms and transcription factors that play a role in the regulation of the transcriptional activity of the oocyte dictating its ultimate acquisition of developmental competence. The oocyte genome has evolved specialized transcription machinery to ensure proper activation of gene that is required for oocyte growth and early embryonic development [71].

## 8. Gene expression and the role of the *Cx43* and *HSPA14* genes during the embryonic development

### 8.1. The *HSPA14* gene

Experiments developed by our team show that exposing cumulus oocyte complexes (COCs) to 41°C did not alter the number of embryos that cleaved but reduced significantly the percentage of development in the blastocyst stage [9]. Additionally, the exposure of bovine embryos to heat shock during oocyte maturation leads to embryos with reduced development and induced alterations in protein synthesis and possibly gene expression as early as the two-cell embryos. In addition, cumulus cells are removed before maturation, which affects and reduces the protein synthesis at 42°C exposure oocytes and COCs. On the other hand, the developed oocytes at 39°C created heat shock protein 70 kDa but oocytes exposure to 42°C did not increase synthesis of any of these proteins, which was shown after examining the methionine- and cysteine-labeled proteins—by two-dimensional SDS-PAGE and fluorography. It has also been noticed that the reduction of protein synthesis caused a prominent decrease in the percentage of protein synthesis in oocytes with intact cumulus compared to those bared oocytes. Therefore, the heat shock increases the steady-state amounts of mRNA for the inducible form of heat shock protein 70 (HSP70) in embryos. The *HSPA14* is the most abundant, highly sensitive to culture environment and it is the major inducible heat shock gene important for protecting embryonic cells from cellular stress in bovine. Thus, HSP70 mRNA in 2- and 4-cell embryos was increased by exposure to 42°C. As findings indicate that the experience for embryos may undergo the transcription of heat shock exposure as early in

2-cell stage, the rates of cleavage and early development are reduced by transcription inhibitors, which show the significance of transcription on the earliest period of fetal life development [72]. The expression of the major heat-inducible protein HSP70 protects cells from a self-destruction known as apoptosis. Cells can also respond to stress by adaptive changes that increase their ability to tolerate normally lethal conditions. The expression of *HSPA14* gene would allow to produce and help maintain cellular function by acting as molecular chaperones to stabilize or refold proteins damaged by heat, and by blocking apoptosis by preventing cytochrome c release from mitochondria during the early stages of apoptosis [73]. Moreover, these authors postulated that the *HSPA14* during embryo development with the presence of antibodies to HSP70 significantly decreases progression to the hatched blastocyst stage in murine embryos. Correspondingly, antisense oligonucleotides complementary to HSP70 mRNA had a similar effect on embryo development, which was amplified ninefold by arsenic exposure at a subtoxic dose. Hence, the embryos that did reach the blastocyst stage despite the inhibition of HSP70 expression, with and without the presence of arsenic, were characterized as degenerate with cell death accompanied by membrane blebbing. Thus, suggested that the requirement for HSP70 during embryo development is amplified by exposure to adverse environmental conditions [12]. *In vitro* matured bovine oocytes, 2-cell and 8-cell embryos, and day 9 hatched blastocysts subjected to control and elevated temperature conditions were analyzed by semiquantitative reverse transcription polymerase chain reaction methods for HSP70 mRNA expression. In the cytoplasm of 8-cell embryos that were inferred under control conditions, HSP70 was evenly disseminated in the cytoplasm but appeared as aggregates in some embryos exposed to elevated temperature. The hatched blastocysts show a competence to react to elevated temperature that is shown in the increased distributions noted after heat stress [74]. As described before, several authors pointed out that heat shock protein can be induced by bovine embryos exposed to heat shock. Those bovine embryos, presented to heat shock, also produce other cellular changes to make cell more impervious to a subsequent heat shock. The 2-cell bovine embryos can deliver HSP70 in response to raised temperature although it is sensitive to heat shock. Therefore, to test for induced thermotolerance, 2-cell bovine embryos were first exposed to a mild heat shock 40°C, allowed to recover at 38.5°C and 5% CO<sub>2</sub>. Subsequently, the way in which HSP70 take part in embryonic development was examined at two different temperature, 38.5 and 40°C, culturing embryos with a monoclonal immunoglobulin to the inducible form of HSP70 experiment, a reduction of the proportion of 2-cell embryos ( $p < 0.05$ ) by adding anti-HSP70 to the culture medium. Thus, bovine 2-cell embryos appear incapable of thermotolerance owing to the increased sensitivity of 2-cell embryos to heat shock as compared to embryos at later stages of development. The results also implicated a role for HSP70 in normal development of bovine embryos. Studies developed by Mayer and Bukau [75] showed that bovine 2-cell embryos appear incapable of thermotolerance owing to the increased sensitivity of 2-cell embryos to heat shock as compared to embryos at later stages of development. The results also implicated a role for HSP70 in normal development of bovine embryos. The same authors [75] pointed out that one of the central components of the cell network of molecular chaperones are HSP70 proteins. An abundant array of protein folding processes are assisted by HSP70 proteins in the cell through association transiently of segments of short hydrophobic peptide—within substrate proteins—with their substrate binding area. When substrate binding, it releases cycle which is

driven by exchanging of HSP70 between the high-affinity ADP bound state and low-affinity ATP bound state. Hence, the chaperone activity of HSP70 proteins need hydrolysis and binding if ATP *in vivo* and *in vitro*. Cochaperones of the J-domain proteins family control this ATPase, which make their substrates targeted by nucleotide exchange factor and by HSP70s, decide the lifetime of the HSP70 substrate complex. This chaperone cycle is tweaked by additional cochaperones. The HSP90 and HSP100 among other chaperones couple with the HSP70 cycle for specific tasks. As a different understanding, Kampinga stated that several assortments of cellular internal and external cases of the stress can be distributed as proteotoxic stresses, which can be described as stresses that raise the main fraction of proteins that are in an unfolded state, thereby promoting the possibility of the formation of intracellular aggregates. These accumulations of assemble, whether not disposed, may lead to the apoptosis. In response to the appearance of denaturation and damaged proteins, cells promote the expression of heat shock proteins. These can have fundamental function as molecular chaperones to avoid protein from the assembling and to maintain the proteins in a state competent for either refolding or degradation. As *Bos indicus* cows usually have better reproductive performance in tropical and subtropical regions than *Bos taurus* cows, presumably due to their better adaptation to tropical environments, Camargo et al. [76] evaluated the developmental competence and expression of the HSP70 gene in immature oocytes from *Bos taurus* (Holstein) and *Bos indicus* (Gyr) dairy cows raised in a tropical region. The total RNA extracted from Holstein and Gyr oocytes and the HSP70 transcripts was conducted, which was formed by real-time PCR after reverse transcription. Cleavage and blastocyst rates were ultimately greater ( $p < 0.05$ ) for Gyr breed ( $n = 390$  oocytes) than for the Holstein ( $n = 505$ ) breed (66.7% vs. 53.1% of cleavage and 19.6% vs. 10.8% of blastocysts, respectively), showing an adaptation of these two breeds to the environmental conditions. Studies developed by Virenque et al. [77] hypothesized that the high structural and functional conservation of HSP during evolution suggests crucial roles in fertilization, embryo development, and thus, fertility in cattle. Also, some HSP are considered housekeeping genes that are essential for many cell functions, such the HSP40 combined with HSP70 acting as chaperones to protect cells from apoptosis [78]. All HSP isoforms are ATP-dependent molecular chaperones as they protect nascent or denatured proteins from aggregation and assist their folding or refolding into native conformation and regulation of heat shock response.

## 8.2. The Cx43 gene

Connexins (CXs) are a family of transmembrane proteins with molecular masses varying from 26 to 60 kD; Cx43 has a molecular mass of 43 kD. In vertebrates, CXs are the building blocks of gap junction channels, intercellular channels that connect the cytoplasm of two neighboring cells. Wrenzycki et al. [79] examined the presence of the mRNA encoding connexin 43 (Cx43) in bovine embryos derived *in vivo* and *in vitro* and in the bovine embryos through morula and blastocysts grown *in vivo*, the transcription of the Cx43 were disclosed. Conversely, when the early *in vitro* stages from cumulus oocyte complexes (COCs) to morula expressed Cx43, blastocysts and hatched blastocysts cannot have disclosed concentrations of mRNA from this gene as because it is not activated or if embryonic gene expression had been active, it terminated prematurely. The differences in transcription between bovine embryos derived *in vivo*

or *in vitro* indicate that culture conditions affect gene expression. Melton et al. [80] reported that the gap junction is a major form of cell-cell communication and aggregations of intercellular channels composed of connexins (CXs), which are responsible for exchange of low molecular weight (<1200 Da) cytosolic materials. In ovarian tissue of mammals, investigators have established that the presence of Cx43 is to determine the ontogeny of early stages of follicular development of embryos. Additionally, it has been found that the preovulatory LH surge caused a decrease in Cx43 mRNA in granulosa cells of rat ovarian follicles. Therefore, the results reported indicated that increased expression of Cx43 temporally correlates with the activation of follicular development and early differentiation of granulosa cells in follicles of prepubertal pig ovaries. Vozzi et al. [81] indicated that in ovarian follicles, oocyte is provided by cumulus cells with small molecules that license control maturation and growth. The germinal cell gets these nutrients through gap junction channels, present between the oocyte and cumulus cells. Goldberg et al. [82] reported that the gap junctions, composed of proteins from the connexin family, are the only channels that directly connect the cytoplasm of the adjacent cells to allow for the intercellular transfer of small hydrophilic molecules, essential for proper development and health in animals and humans. Kidder and Mhawi [83] showed that the gap junctions are a group of intercellular membrane channels which leads to adjacent cells to participate in small molecules (<1 kDa). Gap junction channels are composed of connexins, a homologous family of more than 20 proteins. For developing of follicles, the gap junction is coupled with the developing oocyte and its besetment of the follicle cells into an essential functional syncytium. Among the gap junctions and cumulus cells includes the majority connexin 43, and this connexin has also been detected utilizing immunoelectron microscopy in a small minority of gap junctions at the surface of the oocyte. The significance of connexin 43 for granulosa cell having fundamental function is demonstrated by the fact that follicles lacking this connexin arrest in early prenatal stages and produce incompetent oocytes. Connexin 37 is shown to be the only connexin participate by oocytes to the gap junctions coupling them with granulosa cells, and loss of this connexin interferes with the development of the secondary follicle. The expression of multiple connexins in growth follicles probably indicate the multiple functions served by gap junctional correspondence in folliculogenesis. In the absence of Cx43, granulosa cells block growing in an early preantral stage. To illustrate the fundamental role function of Cx43, dye injection experiments detected that granulosa cells from Cx43 knockout follicles are not coupled, and this was confirmed by ionic current injections. However, electron microscopy detected that gap junctions are quite rare in mutant granulosa cells. Conversely, mutant granulosa cells were eligible to form gap junctions with wild-type granulosa cells in a dye preloading assay. It was accomplished that mutant granulosa cells comprise a population of connexins, composed of nonrecognized connexins that cannot contribute regularly to participate in gap junctions. Additionally, despite Cx43 being the only gap junction protein existing in granulosa cells of early preantral follicles, it is the only one that makes a significant contribution to intercellular coupling [84]. Research conducted by Veitch et al. [85] uncovered that studies of mammalian animals with targeted disruption of specific connexin genes have revealed that at least two connexins, connexin 37 (Cx37) and connexin 43 (Cx43), play essential roles in ovarian follicle development. Immunofluorescence microscopy located Cx37 within gap junction plaques between granulosa cells and the oocyte, and Cx43 between surrounding granulosa cells. Gittens et al. [86] reviewed that the intercellular

communication is required for ovarian folliculogenesis. Yogo et al. [87] reported that the connexin 43 (Cx43)-intermediate gap junctional communication in granulosa cells is crucial for germ line development and postnatal folliculogenesis. These authors showed that follicle-stimulating hormone (FSH) induced the phosphorylation of Cx43 in primary granulosa cells and then further specified to identify Ser365, Ser368, Ser369, and Ser373 in the carboxy terminal tail as the major sites of phosphorylation by FSH, and found that the phosphorylation of these remains was substantial for channel activity. Furthermore, Borowczyk et al. [88] evaluated the role of gap junctions in the regulation of progesterone secretion. The changes in Cx43 mRNA expression were positively correlated with changes in progesterone concentration. They demonstrated a relationship between gap junctions and progesterone secretion that was supported by (1) the positive correlations between progesterone secretion and Cx43 mRNA expression and gap junctional intercellular communication (GJIC) of luteal cells and (2) the inhibition of Cx43 mRNA expression resulted in decreased production of progesterone by luteal cells. This suggested that gap junctions may be involved in the regulation of steroidogenesis in the corpus luteum. Gershon et al. [89] reported that the gap junctions that allow the direct communication between cytoplasmic compartments of neighboring cells are present in a variety of tissues and organs, allowing thus the rapid exchange of ions and metabolites. Besides in the bovine, follicles express Cx43, which is localized to granulosa cells from the primary follicle and increases upon antrum formation being necessary for follicle development and oocyte growth. The expression specificities of Cx43 in ovary are still in discussion by several groups. During folliculogenesis, the granulosa cells give physical support and mediate signal between different types of follicular cells. Rhett et al. [90] reviewed that the Cx43 is the most ubiquitous connexin, with expression in at least 46 different cell types being a gap junction protein widely expressed in mammalian tissues that mediates cell-to-cell coupling. Intercellular channels comprising GJ aggregates form from docking of paired connexons, with one each contributed by opposing cells. The regulation of gap junction coupling is a necessary component of cellular function and response to physiological and pathological stimuli. The cells can modify the expression, phosphorylation state, and protein interactions of Cx43 throughout the cell cycle. Because of the short half-life of Cx43, which is as little as 1–2 h, regulation appears to exist on both short and long time scales through phosphorylation and protein interaction, and gene expression, respectively. Currently, there is a great deal of data on phospho- and transcriptional regulation of Cx43 but less is known about the mechanistic basis and function of Cx43 protein–protein interactions. Furthermore, concerning the Cx43, normally low expression of Cx43 in different development stages of embryos has been associated with low quality and reduced survival capacity of embryos [91, 92].

## 9. Methods for studying differential gene expression

Several techniques have been developed for the screening of genomic alterations at the mRNA level, including subtractive hybridization, differential display-PCR, expressed sequence tag (EST), serial analysis of gene expression (SAGE), and microarray hybridization. Some of these techniques have been used to investigate changes in gene expression at oocyte cell

development and differentiation [93]. In addition, the effective and simple methods for identifying and isolating those genes that are differentially expressed in various cells or under altered conditions. Advantages of the technique include the ability to isolate genes with no prior knowledge of their sequence or identity and the use of common molecular biology techniques that do not require specialized equipment or analyses, its abilities to compare multiple experimental samples for any number of treatment conditions, phenotypes, or genotypes simultaneously, and to identify genes that are either up- or downregulated in one sample relative to another [94]. The differential display include as first step is reverse transcription of mRNA to cDNA using one of anchor primers (which is usually poly T oligonucleotide with one or two additional bases, e.g., T<sub>12</sub> AC) designed to anneal to the 3' poly-A tail of messenger RNA (mRNA). The resulting of cDNA species is subsequently used as templates in a PCR, utilizing the same anchor primer from reverse transcription (RT) reaction in combination with an arbitrary primer. The PCR products may be labeled by incorporation of a radiolabeled nucleotide fluorescently labeled primer. Electrophoresis pattern can be then evaluated by comparing the relative intensities of bands produced from different treatment samples. Bands that are present in one sample and absent in another or bands that are present at different relative intensities, across different treatment, represent potentially differentially expressed mRNA transcripts. The final phase of differential display is excising and purification of the band of interest from the polyacrylamide gel and reamplification by PCR, and eventually subjected to sequencing.

## Acknowledgements

The authors thank Dr. Sofia Quadros and Mr. Filipe António for their precious collaboration in this review. CITA-A is also fully acknowledged. The second author is partially supported by the Azorean Agency for Science and Technology, grant BD M3.1.2/F/044/2011.

## Author details

Aya K. H. Mahdy, Krishna C. Pavani, Erica Baron and Fernando Moreira da Silva\*

\*Address all correspondence to: jsilva@uac.pt

Department of Animal Reproduction, Faculty of Agrarian Sciences and Environment, CITA-A (Research Centre for Agricultural and Environmental Sciences and Technology of the Azores), University of the Azores, Angra do Heroísmo, Portugal

## References

- [1] Sherwin, A., Montgomery, I. and Lundy, G. The impact and implications of climate change for bats. *Mammal Rev.* 2012;43:1–18. DOI: 10.1111/j.1365-2907.2012.00214.x.

- [2] Luo, J., Koselj, K., Zsebok, S., Siemers, M. and Goerlitz, R. Global warming alters sound transmission: differential impact on the prey detection ability of echolocating bats. *J R Soc Interface*. 2013;11(91):20130–20961. DOI: 10.1098/rsif.2013.0961.
- [3] Vermaa, K., Prasadb, S., Mohantyb, T., Kumaresanb, A., Layekb, S., Patbandhac, T., Dattad, T. and Chandb, S. Effect of short-term cooling on core body temperature, plasma cortisol and conception rate in Murrah buffalo heifers during hot-humid season. *J Appl Anim Res.* 2016;44(1):281–286. DOI: 10.1080/09712119.2015.1031782.
- [4] Badinga, L., Thatcher, W., Diaz, T., Drost, M. and Wolfenson, D. Effects of environmental heat stress on follicular development and steroidogenesis in lactating Holstein cows. *Theriogenology*. 1993;39:797–810. DOI: 10.1016/0093-691X(93)90419-6.
- [5] Al-Katanani, M., Paula-Lopes, F. and Hansen, J. Effect of season and exposure to heat stress on oocyte competence in Holstein cows. *J Dairy Sci.* 2002;85(2):390–396. DOI: 10.3168/jds.S0022-0302(02)74086-1.
- [6] Al-Katanani, M. and Hansen, J. Induced thermotolerance in bovine two cell embryos and the role of heat shock protein 70 in embryonic development. *Mol Reprod Dev.* 2002;62(2):174–180. DOI: 10.1002/mrd.10122.
- [7] Ahmed, A., Tiwari, R., Mishra, G., Jena, B., Dar, M. and Bhat, A. Effect of environmental heat stress on reproduction performance of dairy cows. A review. *Int J Livest Res.* 2015;5(4):10–18. DOI: 10.5455/ijlr.20150421122704.
- [8] Roth, Z. Physiology and Endocrinology Symposium: cellular and molecular mechanisms of heat stress related to bovine ovarian function. *J Anim Sci.* 2015;93(5):2034–2044. DOI: 10.2527/jas.2014-8625.
- [9] Pavani, K., Carvalhais, L., Faheem, M., Chaveiro, A., Reis, F. and da Silva, F. Reproductive performance of Holstein dairy cows grazing in dry-summer subtropical. *Asian-Australas J Anim Sci.* 2015;28(3):334–342. DOI: 10.5713/ajas.14.0480.
- [10] Nandi, S., Raghu, M., Ravindranatha, M. and Chauhan, S. Production of buffalo (*Bubalus bubalis*) embryos in vitro: premises and promises. *Reprod Domest Anim.* 2002;37(2):65–74. DOI: 10.1046/j.1439-0531.2002.00340.x.
- [11] Wrenzycki, C., Herrmann, D. and Niemann, H. Messenger RNA in oocytes and embryos in relation to embryo viability. *Theriogenology*. 2007;68(1):77–83. DOI: 10.1016/j.theriogenology.2007.04.028.
- [12] Matwee, C., Kamaruddin, M., Betts, H., Basrur, K. and King, A. The effects of antibodies to heat shock protein 70 in fertilization and embryo development. *Mol Hum Reprod.* 2001;7(9):829–837. DOI: 10.1093/molehr/7.9.829.
- [13] Cushman, A. Physiology and Endocrinology Symposium: the current status of heat shock in early embryonic survival and reproductive efficiency. *J Anim Sci.* 2013;91(3):1141–1142. DOI: 10.2527/jas2013-6231.

- [14] Miova, B., Dinevska-Kjovkarovska, S., Esplugues, V. and Apostolova, N. Heat stress induces extended plateau of hsp70 accumulation-a possible cytoprotection mechanism in hepatic cells. *J Cell Biochem.* 2015;**116**(10):2365–2374. DOI: 10.1002/jcb.25187.
- [15] Li, H., Pang, Q., Qin, L., Jin, S., Zeng, X., Bai, Y. and Li, W. HSP70 overexpression may play a protective role in the mouse embryos stimulated by CUMS. *Reprod Biol Endocrinol.* 2015;**16**(13):125. DOI: 10.1186/s12958-015-0123-z.
- [16] Sharma, T., Nath, A., Prasad, S., Singhal, S., Singh, N., Gade, E., Dubey, K. and Saikumar, G. Expression and characterization of constitutive heat shock protein 70.1 (HSPA-1A) gene in in vitro produced and in vivo- derived buffalo (*Bubalus bubalis*) embryos. *Reprod Domest Anim.* 2012;**47**(6):975–983. DOI: 10.1111/j.1439-0531.2012.02002.x.
- [17] Kidder, G. Roles of gap junctions in ovarian folliculogenesis: implications for female infertility. In: "Gap Junctions in Development and Disease." Springer-Verlag Berlin Heidelberg. 2005, pp. 223–237. DOI: 10.1007/3-540-28621-7\_10.
- [18] Eppig, J. Intercommunication between mammalian oocytes and companion follicle cells. *Bioessays.* 1991;**12**:569–574. DOI: 10.1002/bies.950131105.
- [19] Pointis, G., Fiorini, C., Defamie, N. and Segretain, D. Gap junctional communication in the male reproductive system. *Biochim Biophys Acta.* 2005;**1719**(1–2):1–16. DOI: 10.1016/j.bbamem.2005.09.017.
- [20] McGee, A. and Hsueh, J. Initial and cyclic recruitment of ovarian follicles. *Endocr Rev.* 2000;**21**(2):200–214. DOI: 10.1210/edrv.21.2.0394.
- [21] Senbon, S., Hirao, Y. and Miyano, T. Interactions between the oocyte and surrounding somatic cells in follicular development: lessons from in vitro culture. *J Reprod Dev.* 2003;**49**:259–269. DOI: 10.1262/jrd.49.259.
- [22] De Felici, M., Klinger, G., Farini, D., Scaldaferrri, L., Iona, S. and Lobascio, M. Establishment of oocyte population in the foetal ovary: primordial germ cell proliferation and oocyte programmed cell death. *Reprod Biomed Online.* 2005;**10**(2):182–191. DOI: 10.1016/S1432-6483(10)60939-X.
- [23] Brunet, S. and Verlhac, H. Positioning to get out of meiosis: the asymmetry of division. *Hum Reprod Update.* 2011;**17**(1):68–75. DOI: 10.1093/humupd/dmq044.
- [24] Fair, T. Follicular oocyte growth and acquisition of developmental competence. *Anim Reprod Sci.* 2003;**78**:203–216. DOI: 10.1016/S0378-4320(03)00091-5.
- [25] Barnett, R., Schilling, C., Greenfeld, R., Tomic, D., and Flaws, A. Ovarian follicle development and transgenic mouse models. *Hum Reprod Update.* 2006;**12**(5):537–555. DOI: 10.1093/humupd/dml022.
- [26] Mermilliod, P., Dalbies-Tran, R., Uzbekova, S., Thélie, A., Traverso, M., Perreau, C., Papillier, P. and Monget, P. Factors affecting oocyte quality: who is driving the follicle? *Reprod Domest Anim.* 2008;**43**(2):393–400. DOI: 10.1111/j.1439-0531.2008.01190.x.

- [27] Young, M. and McNeilly, S. Theca: the forgotten cell of the ovarian follicle. *Reproduction*. 2010;140(4):489–504. DOI: 10.1530/REP-10-0094.
- [28] Georges, A., Auguste, A., Bessière, L., Vanet, A., Todeschini, L. and Veitia, A. FOXL2: a central transcription factor of the ovary. *J Mol Endocrinol*. 2013;52(1):17–33. DOI: 10.1530/JME-13-0159.
- [29] Hansen, P. Antecedents of mammalian fertility: lessons from the heat-stressed cow regarding the importance of oocyte competence for fertilization and embryonic development. *Animal Front*. 2014;3(4):34–39. DOI: 10.2527/af.2013-0031.
- [30] Kakkassery, M., Vijayakumaran, V. and Sreekumaran, T. Effect of cumulus oocyte complex morphology on in vitro maturation of bovine oocytes. *J. Vet. Anim. Sci.* 2010;41:12–17.
- [31] Stricker, A., Escalona, R., Abernathy, S. and Marquardt, A. Pharmacological analyses of protein kinases regulating egg maturation in marine nemertean worms: a review and comparison with Mammalian eggs. *Mar Drugs*. 2010;8(8):2417–2434. DOI: 10.3390/ md8082417.
- [32] Homer, H. New insights into the genetic regulation of homologue disjunction in mammalian oocytes. *Cytogenet Genome Res*. 2011;133(2–4):208–222. DOI: 10.1159-000324118.
- [33] Kwon, S. and Lim, J. Small GTPases and formins in mammalian oocyte maturation: cytoskeletal organizers. *Clin Exp Reprod Med*. 2011;38(1):1–5. DOI: 10.5653/cerm.2011.38.1.1.
- [34] Blanco, R., Demyda, S., Moreno, M. and Genero, E. Developmental competence of in vivo and in vitro matured oocytes: a review. *Biotechnol Mol Biol Rev*. 2011;6(7):155–165.
- [35] Da Broi, G., Malvezzi, H., Paz, C., Ferriani, A. and Navarro, A. Follicular fluid from infertile women with mild endometriosis may compromise the meiotic spindles of bovine metaphase II oocytes. *Hum Reprod*. 2014;29(2):315–323. DOI: 10.1093/humrep/det378.
- [36] Dumesic, A., Meldrum, R., Katz-Jaffe, G., Krisher, L. and Schoolcraft, B. Oocyte environment: follicular fluid and cumulus cells are critical for oocyte health. *Fertil Steril*. 2015;103(2):303–316. DOI: 10.1016/j.fertnstert.2014.11.015.
- [37] Ferreira, M., Vireque, A., Adona, R., Meirelles, V., Ferriani, A. and Navarro, A. Cytoplasmic maturation of bovine oocytes: structural and biochemical modifications and acquisition of developmental competence. *Theriogenology*. 2009;71(5):836–848. DOI: 10.1016/j.therionology.
- [38] Wang, Y., Wang, H., Zou, Y. and Xu, M. Mitochondrial functions on oocytes and pre-implantation embryos. *J Zhejiang Univ Sci B*. 2009;10(7):483–492. DOI: 10.1631/jzus. B0820379.
- [39] Kalo, D. and Roth, Z. Effects of mono (2-ethylhexyl) phthalate on cytoplasmic maturation of oocytes – the bovine model. *Reprod Toxicol* 2015;53:141–151. DOI: 10.1016/j.reprotox.2015.04.007.

- [40] Sirard, A. Resumption of meiosis: mechanism involved in meiotic progression and its relation with developmental competence. *Theriogenology*. 2001;55(6):1241–1254. DOI: 10.1016/S0093-691X(01)00480-0 PMID: 11327682.
- [41] Eichenlaub-Ritter, U. and Peschke, M. Expression in in-vivo and in-vitro growing and maturing oocytes: focus on regulation of expression at the translational level. *Hum Reprod Update*. 2002;8(1):21–41. DOI: 10.1093/humupd/8.1.21 PMID: 11866238.
- [42] Tomek, W., Torner, H. and Kanitz, W. Comparative analysis of protein synthesis, transcription and cytoplasmic polyadenylation of mRNA during maturation of bovine oocytes in vitro. *Reprod Domest Anim*. 2002;37(2):86–91. DOI: 10.1046/j.1439-0531.2002.00336.
- [43] Sirard, A., Richard, F., Blondin, P. and Robert, C. Contribution of the oocyte to embryo quality. *Theriogenology*. 2006;65(1):126–136. DOI: 10.1016/j.theriogenology.2005.09.020.
- [44] Niemann, H., Carnwath, W. and Kues, W. Application of DNA array technology to mammalian embryos. *Theriogenology*. 2007;1(68 Suppl 1):S165–S177. DOI: 10.1016/j.theriogenology.2007.05.041.
- [45] Demirtas, E., Holzer, H., Son, W., Chain, R. and Tan, S. In-vitro maturation of human oocytes. In: "Reproductive Endocrinology and Infertility" (Carrell, E. and Peterson, C.). Springer New York. 2010. vol. 5, pp. 633–646. DOI: 10.1007/978-1-4419-1436-1\_42.
- [46] Gasparrini, B. In vitro embryo production in buffalo species: state of the art. *Theriogenology*. 2002;1:237–256. DOI: doi:10.1016/S0093-691X(01)00669-0.
- [47] Ramos, F., Rumpf, R., Câmara, U., Mollo, R., Pivato, I., Marques, R. and Sartori, R. Effect of follicular wave synchronization on in vitro embryo production in heifers. *Anim Reprod Sci*. 2010;117(3–4):201–207. DOI: 10.1016/j.anireprosci.2009.04.009.
- [48] Kathy, N., Niakan, H., Pedersen, R., Simo, C., Renee, A. and Pera, R. Human preimplantation embryo development. *Development*. 2012;139(5):829–841. DOI: 10.1242/dev.060426.
- [49] Wolfenson, D., Roth, Z. and Median, R. Impaired reproduction in heat-stressed cattle: basic and applied aspects. *Anim Reprod Sci*. 2000;61:535–547. DOI: 10.1016/S0378-4320(00)00102-03.
- [50] Rensis, F. and Scaramuzzi, R. Heat stress and seasonal effect on reproduction in the dairy cow – a review. *Theriogenology*. 2003;60:1139–1151. DOI: 10.1016/S0093-691X(03)00126-2.
- [51] Payton, R., Romar, R., Coy, P., Saxton, A., Lawrence, J. and Edwards, J. Susceptibility of bovine germinal vesicle-stage oocytes from antral follicles to direct effects of heat stress in vitro. *Biol Reprod* 2004;71:1303–1308. DOI: 10.1093/biolreprod.104.029892.
- [52] Hansen, J. Exploitation of genetic and physiological determinants of embryonic resistance to elevated temperature to improve embryonic survival in dairy cattle during heat stress. *Theriogenology*. 2007;68:242–249. DOI: 10.1016/j.theriogenology.2007.04.008.
- [53] Roth, Z. Heat stress, the follicle, and its enclosed oocyte: mechanisms and potential strategies to improve fertility in dairy cows. *Reprod Domest Anim*. 2008;2:238–244. DOI: 10.1111/j.1439-0531.2008.01168.x.

- [54] Hansen, J. Stress in dairy animals | heat stress: effects on reproduction. *Phil Trans R Soc Lond B Biol Sci.* 2011;90:3341–3350. DOI: 10.1098/rstb.2009.0131.
- [55] Belhadj, I., Najar, T., Ghram, A. and Abdrrabba, M. Heat stress effects on livestock: molecular, cellular and metabolic aspects, a review. *J Anim Physiol Anim Nutr (Berl).* 2015;100(3):401–412. DOI: 10.1111/jpn.12379.
- [56] Nguyen, T., Bowman, J., Haile-Mariam, M., Pryce, E. and Hayes, J. Genomic selection for tolerance to heat stress in Australian dairy cattle. *J Dairy Sci.* 2016;99(4):2849–2862. DOI: 10.3168/jds.2015-9685.
- [57] Dash, S., Chakravarty, K., Singh, A., Upadhyay, A., Singh, M. and Yousuf, S. Effect of heat stress on reproductive performances of dairy cattle and buffaloes: a review. *Vet World.* 2016;9(3):235–244. DOI: 10.14202/vetworld.2016.235-244.
- [58] Hansen, P. Effects of heat stress on mammalian reproduction. *Phil Trans R Soc Lond B Biol Sci.* 2009;27-364(1534):3341–3350. DOI: 10.1098/rstb.2009.0131.
- [59] Paes, M., Vieira, A., Correia, H., Sa, A., Moura, A., Sales, D., Rodrigues, P., Magalhaes-Padilha, M., Santos, W., Apgar, A., Campello, C., Camargo, S. and Figueiredo, R. Effect of heat stress on the survival and development of in vitro cultured bovine preantral follicles and on in vitro maturation of cumulus-oocyte complex. *Theriogenology.* 2016;86(4):994–1003. DOI: 10.1016/j.theriogenology.2016.03.027.
- [60] Waiz, A., Raies-Ul-Haq, M., Dhanda, S., Kumar, A., Goud, S., Chauhan, S. and Upadhyay, C. Heat stress and antioxidant enzyme activity in bubaline (*Bubalus bubalis*) oocytes during in vitro maturation. *Int J Biometeorol.* 2016;60(9):1357–66. DOI:10.1007/s00484-015-1129-0.
- [61] Jones, T. Turning it on and off: M-phase promoting factor during meiotic maturation and fertilization. *Mol Hum Reprod.* 2004;10(1):1–5. DOI: 10.1093/molehr/gah009.
- [62] Brunet, S. and Maro, B. Cytoskeleton and cell cycle control during meiotic maturation of the mouse oocyte: integrating time and space. *Reproduction.* 2005;130(6):801–811. DOI: 10.1530/rep.1.00364.
- [63] Haccard, O. and Jessus, C. Oocyte maturation, Mos and cyclins--a matter of synthesis: two functionally redundant ways to induce meiotic maturation. *Cell Cycle.* 2006;5(11):1152–1159. DOI: 10.4161/cc.5.11.2800.
- [64] Dumollard, R., Levasseur, M., Hebras, C., Huitorel, P., Carroll, M., Chambon, P. and McDougall, A. Mos limits the number of meiotic divisions in urochordate eggs. *Development.* 2011;138(5):885–895. DOI: 10.1242/dev.057133.
- [65] Zhang, Y., Duan, X., Cao, R., Liu, L., Cui, S., Kim, H., Rui, R. and Sun, C. Small GTPase RhoA regulates cytoskeleton dynamics during porcine oocyte maturation and early embryo development. *Cell Cycle.* 2014;13(21):3390–3403. DOI: 10.4161/cc.15384101.2014.9529.
- [66] Coticchio, G., Dal, M., Mignini, M., Guglielmo, C., Brambillasca, F., Turchi, D., Novara, V. and Fadini, R. Oocyte maturation: gamete-somatic cells interactions, meiotic resumption, cytoskeletal dynamics and cytoplasmic reorganization. *Hum Reprod Update.* 2015;21(4):427–454. DOI: 10.1093/humupd/dmv011.

- [67] Park, T., Seo, M., Lee, Y. and Lee, A. Lin28 regulates the expression of neuropeptide Y receptors and oocyte-specific homeobox genes in mouse embryonic stem cells. *Clin Exp Reprod Med.* 2012;39(2):87–93. DOI: 10.5653/cepm.2012.39.2.87.
- [68] Rajkovic, A. and Matzuk, M. Functional analysis of oocyte-expressed genes using transgenic models. *Mol Cell Endocrinol.* 2002;221:87(1-2):5–9. DOI: 10.1016/S0303-7207(01)00710-9.
- [69] Rajkovic, A., Yan, C., Yan, W., Klysik, M. and Matzuk, M. Obox, a family of homeobox genes preferentially expressed in germ cells. *Genomics.* 2002;79(5):711–717. DOI: 10.1006/geno.2002.6759.
- [70] Kim, M., Ahn, J., Lee, S., Lee, A., Lee, M., Kim, H., Kim, S. and Park, S. Obox4 regulates the expression of histone family genes and promotes differentiation of mouse embryonic stem cells. *FEBS Lett.* 2010;584(3):605–611. DOI: 10.1016/j.febslet.
- [71] DeJong, J. Basic mechanisms for the control of germ cell gene expression. *Gene.* 2006;366(1):39–50. DOI: 10.1016/j.gene.2005.10.012.
- [72] Chandolia, K., Peltier, R., Tian, W. and Hansen, J. Transcriptional control of development, protein synthesis, and heat-induced heat shock protein 70 synthesis in 2-cell bovine embryos. *Biol Reprod.* 1999;61(6):1644–1648. DOI: 10.1095/biolreprod61.6.1644.
- [73] Mosser, D., Caron, W., Bourget, L., Meriin, B., Sherman, Y., Morimoto, I. and Massie, B. The chaperone function of hsp70 is required for protection against stress-induced apoptosis. *Mol Cell Biol.* 2000;19:7146–7159. DOI: 10.1128/MCB.20.19.7146\_7159.2000.
- [74] Kawarsky, J. and King, A. Expression and localisation of heat shock protein 70 in cultured bovine oocytes and embryos. *Zygote.* 2001;9(1):39–50. DOI: 10.1017/S0967199401001058.
- [75] Mayer, P. and Bukau, B. Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci.* 2005;62(6):670–684. DOI: 10.1007/s00018-004-4464-6.
- [76] Camargo, S., Viana, H., Ramos, A., Serapião, V., de Sa, F., Ferreira, M., Guimarães, F. and do Vale, R. Developmental competence and expression of the Hsp70.1 gene in oocytes obtained from Bos indicus and Bos taurus dairy cows in a tropical environment. *Theriogenology.* 2007;68(4):626–632. DOI: 10.1016/j.theriogenology.2007.03.029.
- [77] Virenque, A., Camargo, S., Serapião, V., Rosa, A., Watanabe, F., Ferreira, M., Navarro, A., Martins, P. and Ferriani, A. Preimplantation development and expression of Hsp-70 and Bax genes in bovine blastocysts derived from oocytes matured in alpha MEM supplemented with growth factors and synthetic macromolecules. *Theriogenology.* 2009;71(4):620–627. DOI: 10.1016/j.theriogenology.2008.09.028.
- [78] Zhang, B., Peñagaricano, F., Driver, A., Chen, H. and Khatib, H. Differential expression of heat shock protein genes and their splice variants in bovine preimplantation embryos. *J Dairy Sci.* 2011;94(8):4174–4182. DOI: 10.3168/jds.2010-4137.
- [79] Wrenzycki, C., Herrmann, D., Carnwath, W. and Niemann, H. Expression of the gap junction gene connexin43 (Cx43) in preimplantation bovine embryos derived in vitro or in vivo. *J Reprod Fertil.* 1996;108(1):17–24. DOI: 10.1016/0093-691X(96)84724-8.

- [80] Melton, M., Zaunbrecher, M., Yoshizaki, G., Patiño, R., Whisnant, S., Rendon, A. and Lee, H. Expression of connexin 43 mRNA and protein in developing follicles of prepubertal porcine ovaries. *Comp Biochem Physiol B Biochem Mol Biol.* 2001;130(1):43–55. DOI: 10.1016/S1096-4959(01)00403-1.
- [81] Vozzi, C., Formenton, A., Chanson, A., Senn, A., Sahli, R., Shaw, P., Nicod, P., Germond, M. and Haefliger, A. Involvement of connexin 43 in meiotic maturation of bovine oocytes. *Reproduction.* 2001;122(4):619–628. DOI: 10.1530/rep.1.004.
- [82] Goldberg, S., Moreno, P. and Lampe, D. Gap junctions between cells expressing connexin 43 or 32 show inverse sperm selectivity to adenosine and ATP. *J Biol Chem.* 2002;277(39):36725–36730. DOI: 10.1074/jbc.M109797200.
- [83] Kidder, M. and Mhawi, A. Gap junctions and ovarian folliculogenesis. *Reproduction.* 2002;123(5):613–620. DOI: 10.1530/rep.0.1230613.
- [84] Gittens, E., Mhawi, A., Lidington, D., Ouellette, Y. and Kidder, M. Functional analysis of gap junctions in ovarian granulosa cells: distinct role for connexin 43 in early stages of folliculogenesis. *Am J Physiol Cell Physiol.* 2003;284(4):880–887. DOI: 10.1152/ajpcell.00277.2002.
- [85] Veitch, I., Gittens, E., Shao, Q., Laird, W. and Kidder, M. Selective assembly of connexin37 into heterocellular gap junctions at the oocyte/granulosa cell interface. *J Cell Sci.* 2004;117(13):2699–2707. DOI: 10.1242/jcs.01124.
- [86] Gittens, E., Barr, J., Vanderhyden, C. and Kidder, M. Interplay between paracrine signaling and gap junctional communication in ovarian follicles. *J Cell Sci.* 2005;118(1):113–122. DOI: 10.1242/jcs.01587.
- [87] Yogo, K., Ogawa, T., Akiyama, M., Ishida-Kitagawa, N., Sasada, H., Sato, E. and Takeya, T. PKA implicated in the phosphorylation of Cx43 induced by stimulation with FSH in rat granulosa cells. *J Reprod Dev.* 2006;52(3):321–328. DOI: 10.1677/JOE-06-0076 PMID: 16474210.
- [88] Borowczyk, E., Johnson, L., Bilski, J., Bilska, A., Redmer, A., Reynolds, P. and Grazul-Bilska, T. Role of gap junctions in regulation of progesterone secretion by ovine luteal cells in vitro. *Reproduction.* 2007;133(3):641–651. DOI: 10.1530/REP-06-0178.
- [89] Gershon, E., Plaks, V. and Dekel, N. Gap junctions in the ovary: expression, localization and function. *Mol Cell Endocrinol.* 2008;282(1–2):18–25. DOI: 10.1016/j.mce.2007.11.001.
- [90] Rhett, M., Jourdan, J. and Gourdie, G. Connexin 43 connexon to gap junction transition is regulated by zonula occludens-1. *Mol Biol Cell.* 2011;22(9):1516–1528. DOI: 10.1091/mbc.E10-06-0548.
- [91] Lonergan, P., Rizos, D., Gutierrez-Adan, A., Moreira, M., Pintado, B., de la Fuente, J. and Boland, P. Temporal divergence in the pattern of messenger RNA expression in bovine embryos cultured from the zygote to blastocyst stage in vitro or in vivo. *Biol Reprod.* 2003;69:1424–1431. DOI: 10.1095/biolreprod.103.018168.

- [92] Rizos, D., Gutierrez-Adan, A., Perez-Garnelo, S., De La Fuente, J. and Boland, L. P. Bovine embryo culture in the presence of serum: implication for blastocyst development, cryotolerance, and messenger RNA expression. *Biol Reprod.* 2003;68:236–243. DOI: 10.1095/biolreprod.102.007799.
- [93] Hamatani, T., Yamada, M., Akutsu, H., Kuji, N., Mochimaru, Y., Takano, M., Toyoda, M., Miyado, K., Umezawa, A. and Yoshimura, Y. What can we learn from gene expression profiling of mouse oocytes?. *Reproduction.* 2008;135(5):581–592. DOI: 10.1530/REP-07-0430.
- [94] Liang, P. and Pardee, B. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science.* 1992;257(5072):967–971. DOI: 10.1126/science.1354393.

---

# **Cells: Segway for Genetically Desired Cattle through Embryotic Development Sequencings and Genetic Mapping**

---

Marleen C. Dinis and Joaquim F. Moreira da Silva

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/65642>

---

## **Abstract**

Cells have a direct relation to understanding cattle genetic information aside from the cell's main role of housing genetic material. Gene presence being visible through phenotypic characteristics of living organisms allows for researchers to pinpoint specific genetic markers to access and evaluate the genetic information for desired visible traits. On a molecular level, through the interaction between cattle embryos and bovine cells, researchers can pair this information with current technological innovation to better understand on how genetics are manipulated naturally by environmental elements, which can have a direct effect on countries where cattle milk and meat production are the major contributors to economic stability. This overview of past, current, and potential future research involving cells can provide an understanding of the importance in cells and their relationship with biotechnology innovation. Through techniques involving cells and genetic analysis, various methods can be utilized to overcome limitations of cattle reproduction and increase the presence of desirable cattle traits, which in turn not only aid in the economic success of countries dependent on cattle products, but can also open the door to human therapeutics.

**Keywords:** genetics, cattle, cell, granulosa, embryo

---

## **1. Introduction**

Since the age of primitive man and of the hunter-gatherer, cattle has been the source of nutrition, clothing, culture, and innovation of tools to improve man's existence. As seen in **Figure 1**, an early view of animals through the eyes of cave dwellers in prehistory time allows today's modern

citizens to witness the existence of cattle (such as the ox) not only as a creatures of coexistence, but also as means of survival. As early as 8000 BC, the bovine species was domesticated as a means of a secure meat supply before becoming an essential object for clothes, tents, and even later becoming the main component in creating drums. Throughout history, cattles have been valuable not only as means of self-survival, but also played a major role in the art of agriculture and the conception of community. As the idea of growing food sources in one area (farming) grew more and more attainable the lifestyle of hunters and gathers phased into established communities with the help of cattle. The use of cattle as work tools for plowing fresh soil, as depicted in **Figure 2**, and the importance of manure for crop fertilization allowed the human race to settle and create civilized communities.



**Figure 1.** Untitled painting of ancient cave art. Retrieved from [https://en.wikipedia.org/wiki/Bison\\_priscus](https://en.wikipedia.org/wiki/Bison_priscus).



**Figure 2.** Untitled Egyptian art. Retrieved from <https://ucfant3145f09-04.wikispaces.com/Ancient+Egyptian+Food+%26+Agriculture>.

Today cattle remains a vital part in mankind's socio and economic progression through the same deliverables attained by their ancestors: food, clothing, and innovative tools. What worries the majority of cattle owners today is the pressure of supply and demand of meat and dairy products, which comes from sustaining high populations of cattle. Past and current

techniques for increasing cattle population include nutritional management and health practices to maintain or increase cattle production. As technological advancements and understanding of cattle genetics progress, methods including cell activity, genetic practices, and bioinformatics tools have been utilized to increase a deeper understanding of the cattle genetics not only to increase livestock numbers, which increase the production of desirable cattle that yields high economic profit, but also open new opportunities for novel human therapeutics.

## 2. Breeding

One of the oldest methods for desirable cattle production involves breeding both close breeding and outbreeding (Figure 3). A popular and effective method of outbreeding, known as crossbreeding, involves breeding animals of pure breeds within each generation to produce desired animals. There are two main categories that fall under the crossbreeding umbrella: terminal- and rotational-crossing. Terminal-crossing is utilized more for meat animals and not dairy animals as it does not involve the progeny of breeding, but rather is slaughtered for commercialized meat. Whereas rotational-crossing involves the breeding of two breeds and utilizing the female offspring to breed with a sire of a different breed. The female offspring of this breeding sequence is introduced to one of the two breeds used in the original breeding scheme [1]. These techniques are particularly important when seeking heterosis. Heterosis is the improvement or enhancement of a specific trait that involves three types of improvement with respect to fertility: individual affecting the calf, maternal affecting the dam, and paternal affecting the sires. Individual heterosis increases live births and postnatal calf survival, while maternal increases conception rate, birth rates, and other economic traits such as milk product [2].

### Close breeding

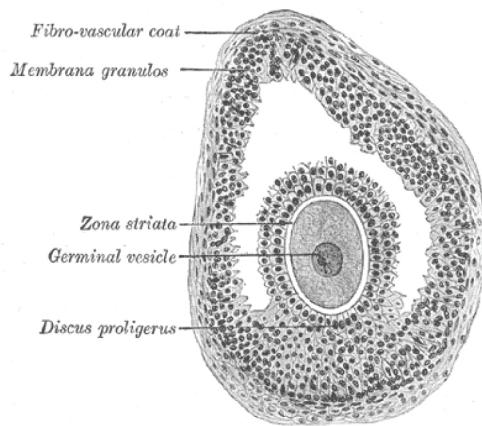
- Inbreeding
- Line breeding

### Outbreeding

- Crossbreeding
- Outcrossing
- Backcrossing
- Topcrossing
- Granding up
- Mating likes
- Mating unlikes

Figure 3. Types of cattle breeding methods.

These techniques are important in producing genetically desirable cattle to improve economic stimulating traits, but novel techniques are being studied to further the progression and overcome variables in increasing cattle. As previously stated nutritional management has been a means of sustaining current cattle numbers, but what if the nutrition in itself becomes an issue of calf abortions and still borns? This is an issue that is not new to those who raise cattle in the archipelago of the Azores. The Azores are made up of nine volcano-derived islands centered in the Atlantic Ocean roughly 2454 km from the coast of the United States and 1600 km from Portugal [3]. Bovine population and production are a vital source of economic stability for the archipelago with products ranging from butter, milk, and cheese to meat supplies for commercial fast food chains. Unfortunately, the Azores face two debilitating periods within the year where animal product is limited due to lack of grass growth. Cattle owners are forced to supplement their cattle with a native invader plant as an alternative feed, of which contains *Pittosporum undulatum* [4]. The compound *Pittosporum undulatum* can have a negative effect on an animal's immune, endocrine, and reproductive system, thus becoming a problem with cattle production even though it is a means of sustaining the current cattle population. A recent study (2016) performed at the University of the Azores provided an in-depth understanding of how this compound can directly affect embryonic development through coculturing with bovine granulosa cells (Figure 4) [5].



**Figure 4.** Depiction of ovarian follicle. The granulosa cells utilized in the previously discussed section of coculture utilized these cells that form the membrane granulose. Retrieved from: [https://en.wikipedia.org/wiki/Ovarian\\_follicle#/media/File:Gray1164.png](https://en.wikipedia.org/wiki/Ovarian_follicle#/media/File:Gray1164.png).

### 3. Coculturing

The idea behind coculturing cells with embryos allows for researchers to evaluate not only the progression of embryonic development, but also pinpoint interactions between the two entities to establish a better understanding of what cause and effects are taking place. Bovine granulosa

cells play a vital role in *in vitro* maturation (IVM) of coculture with oocytes as well as the development of embryos derived from *in vitro* fertilization (IVF) [6]. Granulosa cells express growth factor kit ligand (KL), which assist in follicular development and directly effects oocytes whenever gonadotropins, FSH- and LH-specific levels are present [7]. Through quantitative RT-PCR and specific primers one can analyze the gene expression level of KL. In the case involving *Pittosporum undulatum* previously mentioned, the study did not utilize RT-PCR, but rather relied on the expression of blocked embryo development. The idea behind the study at the University of the Azores stems from this hopeful possibility of understanding how and when embryonic development is affected when *Pittosporum undulatum* is present, which can have a relationship with KL production of the granulosa cells. The study showed that embryos cocultured with granulosa cells collected from cattle that were fed *Pittosporum undulatum* demonstrated higher embryo blocking when cocultured granulosa cells as opposed to the control group that were not exposed to the same feed (Figures 5 and 6).

Treatment	Nº of ovaries	No. of collected oocytes	No. of oocytes for maturation	Rate of		
				Maturation	Cleavage	Developed Embryo
<b>Experimental</b>	8	136	119	78.27±5.81 <sup>a</sup>	29.92±4.31 <sup>a</sup>	7.30±3.11 <sup>a</sup>
<b>Control</b>	8	186	148	90.46±3.02 <sup>a</sup>	41.86±5.58 <sup>b</sup>	21.88±6.85 <sup>a</sup>

Numbers in the same column with different letters (a, b) differ significantly at  $P < 0.05$ .

Data are expressed in percentage (%) as mean  $\pm$  SEM, standard error of means.

% of developed embryos classified as morula, early blastocyst and blastocyst stages.

**Figure 5.** Rates of development embryonic *in vitro* of experimental and control groups.

### 3.1. Cells and transgenic techniques

Another interesting factor in applying cocultures with cells involves genetics directly by creating transgenic animals. A transgene is a genetic material that is naturally transferred from one organism to another through genetic engineering, which can change the genotype and phenotype of the transgenic organism. This process utilizes newly fertilized cells that are isolated and injected with specific DNA, the egg is then transferred into a host mother where it is able to develop into an embryo. Currently, a new method involves the removal of the nucleus from one egg and replaces with the nucleus of another cell; this is commonly known as nuclear transfer. This allows for complete transfer of genetic information of the donor cell into the recipient. Further yet the cells may be genetically modified with a laboratory before being transferred to the host. This type of cellular genetic transfer was conducted at the University of Georgia (Athens, Georgia) focused on the production of transgenic bovine

embryos through a granulosa cell theater [8]. This allows for the development of genetically desirable cattle through selection of genetic enhancement which can increase specific phenotypes of economically pleasing cattle. The study conducted at the University of Georgia showed that bovine granulosa cells can transfect desirable genes *in vitro* without negatively affecting embryo viability; essentially bovine granulosa cells are potentially useful passage cells to establish nuclear transfer of gene cloning [8]. Transgenic animals can add an advantage to the human therapeutic industry by providing preclinical theatres for vaccine production and tissue growth for human use.



**Figure 6.** Picture of bovine granulosa cells utilized at the University of the Azores. Photograph taken by Dr. Helder Nunes at the University of the Azores.

### 3.2. Cells and cryopreservation advancement

Further evidence suggests that the use of various bovine cell lines can assist in the cryopreservation of bovine embryos. This becomes essential when one applies genetically desirable crossbreeding *in vitro* to create embryo offspring in numbers, especially with the current technique of offspring sex selection. Studies show that cell culture systems with bovine oviduct epithelium cells (BOECs) can be utilized *in vitro* to increase embryo robustness [9]. BOECs are comprised of ciliated and nonciliated (secretory) cells. The ciliated cells aid in moving the ovum away from the ovary and toward the uterus, while the nonciliated cells release secretions needed to lubricate, provide nourishment, and protection for traveling ovum. Furthermore, it has been proven that BOEC can have a direct effect on embryo development when present during the first 4 days or the last 4 days of development by accelerating blastocyst development [10]. Prior to cryopreservation the bovine oocytes are matured and fertilized *in vitro* in culture medium supplemented with serum, as its presence provides a beneficial environment, such as energy substrate, amino acids, vitamins, and growth factors. The serum has a biphasic effect on embryonic development, on the one hand by inhibiting the first division, and on the other, by encouraging the development of blastocysts [11]. Thus, the use of BOECs and its metabolic entities with bovine embryos in simple medium can aid in furthering the embryo environment into a rich maturation media. The use of BOECs in a culture system with embryos

has not been exhausted and has not been studied in the comparison of “fresh” embryos and cryopreserved embryos. Second, according to Jang et al., it is highly recommended that BOECs be studied to investigate the specific mechanism of their actions with the development of embryos as well as provide a complete analysis of the somatic cell coculture system that supports high development rates *in vitro* as well as a potential increase in viability rates *in vivo* in both fresh and frozen embryos [12]. According to Schmaltz-Panneau et al., the mechanisms of BOECs have not been clearly clarified, but evidence shows that *in vitro* coculture of bovine embryos with BOECs can mimic the maternal environment which can improve the developmental quality of the embryo [15].

Previous studies suggest that tissue inhibitor of the enzyme metalloproteinase-1 produced in BOECs possess embryogenesis-stimulating activity, while another study demonstrated that BOECs contained luteinizing hormone (LH) receptors that are functional in increasing the synthesis of oviductal glycoprotein, which can increase the development of early embryos [13]. BOECs also control oxygen levels within the embryo environment, which is a vital factor in embryo success. To assist with embryo and maternal communication oviduct-secreted proteins (OSP) create a pathway for this communication as well as the maintenance of environment for successful embryo development [14]. There has been a progression in identifying many OSP, but two antioxidant enzymes seem to play a major part in continued embryo development: copper, zinc superoxide dismutase (Cu, Zn-SOD), and phospholipid hydroperoxide glutathione peroxidase (GPx-4) [14]. The metabolic exchange between SOD and GPx-4 maybe the explanation for the increase in the blastocyst rate of embryos when BOEC is present due to the gas regulation during development. BOECs also regulate the secretion of embryotrophic factors such as growth factors, component C3, oviduct, and osteopontin. The osteopontin protein is believed to have a direct effect on spermatozoa, oocytes, and embryos [15]. Osteopontin was first identified in the mineralized matrix of bovine bone, but it is also present in the testis and epididymis in males as well as within the female reproductive tissues including the oviduct. Further investigation of osteopontin determined that this protein is capable of enhancing bovine cleavage and blastocyst rates in culture [15]. This information sparks the question if a better genetic understanding of this occurrence is understood could researchers essentially pinpoint and modify this communication between embryo and its maternal surrounding (whether natural or synthetic) to aid in the fight against bovine miscarriages and to progress in successful postcryopreserved embryo development rates once they are introduced into carrying host. One way of using this information regarding cellular interaction is to understand the genetic phenomena that occur during specific phases of cellular expression. Another is to utilize this advancement in embryo cryopreservation in order bank transgenic embryos.

#### 4. Genetic mapping

Improving genetic understanding and innovative techniques have long evolved since the first recorded line of breeding performed by Robert Bakewell and the first freezing of semen in 1949. In 2001, genomic selection and genetic mapping became a popular method in improving

livestock genetics. Identifying and cloning genetically desirable traits can create a robust mapping of bovine genes. Genetic mapping is essential for researchers to identify where genes are located and how they function within cells and overall how it effects the phenotype of cattle. Genetic mapping of humans and cattle alike plays crucial roles in understanding chromosomal rearrangement through evolutionary progression. The understanding of cattle genome can shine a light on evolutionary changes but can also illuminate a path to pinpointing and selecting desirable and marketable phenotypes. Since it can be extremely costly to select specific desirable animals, it is beneficial to start the process as early as possible which makes embryo production of selective cattle not only more cost efficient, but also effective in ensuring that the trait is carried within the genetic information of the embryo.

#### 4.1. Synteny mapping

A common approach to developing a genetic map is that of synteny mapping which utilized somatic cells. Synteny refers to two or more genomic regions within a set of chromosomes within a species that show evidence of deriving from one ancestral genomic region [16]. Synteny mapping consists of a method utilizing a generated list of well conserved genes, known as *syntentic blocks*, which are associated with the same chromosome for a specific species; whereas conserved synteny describes the location of homologous genes on the same chromosome, but for two different species [17]. Somatic cells are the primary component of synteny mapping, and they are living mammalian cells that are not gamete cells (i.e., ova and sperm cell) such as those that make up bones, organs, and blood cells. Theoretically, any gene product or sequence can be successfully mapped using the synteny method as long as the presence (or absence) of the gene of the species in question can be challenged against the fully retained rodent genomic background (through hybrid somatic cells) to confirm whether a concordance or discordance for the location of two genes is located on the same chromosome [17].

#### 4.2. Bioinformatics and genomic selection targets

Today bioinformatics friendly programs such as BLAST are used to identify specific regions with similar genetic sequences for better understanding of characterizing organism and biological discoveries. Specific sites such as bovinegenome.org provide search and annotation tools to support bovine genome and those researchers seeking to expand bovine genetic mapping. Studies such as those conducted by Homer et al. touch upon genetic solutions to the increasing problem with milk production. Poor expression of estrus is considered to be a contributing factor to poor dairy cow fertility; by identifying where the genetic issue lies, Homer and his team believe that they can identify which cows display higher expressions of estrus, as well as identify the genetic issue associated with this lack of expression [18].

From a traditional standpoint, dairy cows are bred based on simple breeding selection, which primarily depends on the sire-side with the aid of worldwide semen distribution through artificial insemination. The use of genomic selection is used in dairy cattle to overcome limitation of traditional selection due to health and fertility problems. Genomic selection for dairy cattle owners becomes beneficial when comparing costs, reproduction, and generation

cycles associated with traditional breeding using progeny testing. There are two ways one can use genetic selection for breeding: one way is by preselecting young bulls to test or the other option involves selecting the bull based on already available genomic information, thus the use of genetic selection for future dairy cow reproduction can have a major positive impact on the dairy industry [19]. As for beef, cattle selection of desired animals is based on the targeted market, and due to the large population of beef cattle and lower accuracy of beef cattle genetic markers, the idea of using genetic selection is not as desirable and economically pleasing as it is for dairy cows [20].

## 5. Conclusion

Conventional breeding involves hundreds of unidentified genes, which require continued advancement in modern molecular biology with gene mapping. By identifying and isolating specific genes to create transgenic animals, as well as with the aid of cells to advance the understanding of environmental influences and their ability to improve cryopreservation of genetically desired embryos. Through better understanding of cellular interaction and genetic traits, researchers seeking answers for issues regarding animal reproduction can ultimately find solutions to not only maintain stable offspring numbers, but also produce genetically desirable progenies.

As mentioned, new phenotype genetic selection along with easier transfer of genetics can answer the social demand of meat and dairy products through preferable phenotype production. In the future, the dairy and meat industry will be faced with higher populations and not enough cattle products to sustain the increasing populace, as such the need for faster reproduction cycles and robust dairy cows will be necessary to meet demands. Future researchers are urged to expand their knowledge of the cattle genetic map, which allows for a better understanding not only of possible genetic selection for breeding purposes, but also aids in understanding the human genome itself. This is especially true when focusing on reproduction, when considering both human female and a cow, the only difference between the two is the length of the estrus cycle. Further studies on genetic relation to estrous expression can not only aid cattle owners in effectively increasing cattle head but also can aid in human infertility-related research as well.

A future capability of cellular participation in innovation is endless. Here a very small number of potential uses of cells for genetic selection are briefly introduced. There are numerous uses of cells to expand our knowledge of veterinary genetics. Even though cells house the information that makes everything what it appears to be, research progression in biotechnology has opened a portal to understanding how these fragments of information function and (to an extent) how to manipulate them to better understand their capabilities in creating their natural end product. A global perspective cells are minuscule in size but hold an enormous amount of information and material for researchers to utilize the information for animal and human therapeutic innovation.

## Acknowledgements

The authors would like to thank the University of the Azores for allowing me to be a scientific collaborator, specifically Professor Joaquim F. Moreira da Silva, PhD and Dr. Helder Patricio Barcelos Nunes for allowing me to help and for introducing us to the amazing capabilities of the granulosa cells.

## Author details

Marleen C. Dinis\* and Joaquim F. Moreira da Silva

\*Address all correspondence to: marleencook@gmail.com

The University of the Azores, Angra do Heroísmo, Azores, Portugal

## References

- [1] FAO. Chapter 6 Crossbreeding [Internet]. 1982 [Updated: 2016]. Available from: <http://www.fao.org/docrep/004/X6536E/X6536E06.htm>
- [2] Greiner, S. Crossbreeding Beef Cattle [Internet]. May 1, 2009. Available from: <https://pubs.ext.vt.edu/400/400-805/400-805.html>
- [3] Visit Azores. The Archipelago [Internet]. 2016. Available from: <http://www.visit-azores.com/en/the-azores/the-9-islands/the-archipelago/geography>
- [4] Nunes, H., Faheem, M., Dinis, M., Borba, A., da Silva, F.M. Effect of feed with *Pittonsporum undulatum* in vivo on bovine progesterone levels and embryo produced in vitro. Canadian Journal of Animal Science. Forthcoming. DOI: CJAS2
- [5] Nunes, H.P., Furnas, S., Dinis, M., Borba, A., Da Silva, F.M. Maturation, fertilisation and culture of bovine oocyte using chlorogenic acid in media. Animal Science. Forthcoming. DOI:
- [6] Pinkert, C.A., editors. A Laboratory Handbook. 3rd ed. Waltham, MA: Elsevier; 2014; p. 696. DOI: 978-0-12-410490-7
- [7] Parrot, J.A., Skinner, M.K. Thecal cell-granulosa cell interactions involved a positive feedback loop among keratinocyte growth factor, hepatocyte growth factor, and kit ligand during ovarian follicular development. Endocrinology. 1998;139(5):2240–2245.
- [8] Sezen, A., Rzucidlo, S.J., Gibbons, J., Kazuchika, M., Stice, S.L. Production of transgenic bovine embryos by transfer of transfected granulosa cells into enucleated oocytes. Molecular Reproduction and Development. 2001;(60):20–26.

- [9] Abe, H., Hoshi, H. Bovine oviductal epithelial cells: Their cell culture and applications in studies for reproductive biology. *Cytotechnology*. 1997;23(1):171–183.
- [10] Cordova, A., Perreau, C., Uzbekova, S., Ponsart, C., Locatelli, Y., Mermilliod, P. Development rate and gene expression of IVP bovine embryos cocultured with bovine oviduct epithelial cells at early or late state of preimplantation development. *Theriogenology*. 2014;81(1):1163–1173.
- [11] Jonge, C., Barratt, C., editors. *Assisted Reproductive Technology*. New York: Cambridge University Press; 2002. 435 p.
- [12] Jang, H.Y., Jung, Y.S., Cheong, H.T., Kim, J.T., Park, C.K., Kong, H.S., et al. Effects of cell status of bovine oviduct epithelial cell (BOEC) on the development of bovine IVM/IVF embryos and gene expression in the BOEC used or not used for the embryo culture. *Asian Australasian Journal of Animal Science*. 2008;21(7):980–987.
- [13] Ian Gordon. *Laboratory Production of Cattle Embryos*. 2nd ed. Cambridge, MA: CABI; 2003. 549 p.
- [14] Mermilliod, P., Schmaltz, B., Perreau, C., Tslkis, G., Martinot, E., Cordova, A., et al. Early development of ruminant embryos, autonomous process or the result of a positive dialog with surrounding maternal tissues. *Acta Scientiae Veterinariae*. 2010;38(2):521–533.
- [15] Schmaltz, B., Cordova, A., Dhorne, S., Hennequet, C., Uzbekova, S., Martinot, E., et al. Early bovine embryos regulate oviduct epithelial cell gene expression during in vitro co culture. *Animal Reproduction Science*. 2014;149:103–116.
- [16] Cogepedia. Synteny [Internet]. July 24, 2014. Available from: <https://genomevolution.org/wiki/index.php/Synteny>
- [17] Womack, J.E. The cattle gene map. *ILAR Journal*. 1998;39(2):153–159.
- [18] Homer, E., Derecka, K., Webb, R., Garnsworthy, P. Mutations in genes involved in oestrous cycle associated expression of oestrus. *Animal Reproduction Science*. 2013;142:106–112.
- [19] Jonas, E., Koning, D. Genomic selection needs to be carefully assessed to meet specific requirements in livestock breeding program. *Frontiers in Genetics*. 2015;6(49):1–8.
- [20] Johnston, D., Tier, B., Graser, H. Beef cattle breeding in Australia with genomics: opportunities and needs. *Animal Production Science*. 2012;52:100–106.



## Beyond Fifty Shades: The Genetics of Horse Colors

---

Adriana Pires Neves, Eduardo Brum Schwengber,  
Fabiola Freire Albrecht, José Victor Isola and  
Liana de Salles van der Linden

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/65848>

---

### Abstract

Since the dawn of horse domestication, coat colors have always fascinated humankind. In the last century, knowledge of genetics and development of scientific tools have become powerful enough so that the effects of many DNA mutations could be critically studied. Coat color nomenclature varies according to countries and breed associations; in addition, many factors can modify the color of the coat, such as sun exposure, age, sex, and nutritional status of the animal. Nevertheless, horses are capable of producing only two pigments. Several genes have been indicated as putative to coat color modification, altering the basic color by dilution, redistribution, or lacking of pigments.

**Keywords:** horse, genetics, coat color, alleles, DNA

---

### 1. Introduction

Nearly 6000 years ago, an extraordinary event would change the course of human history: the domestication of horse, which began in the region known today as Kazakhstan [1], in Central Asia. Initially looking for food and materials for making clothes and tools, the primitive man had also found an ally that would provide relatively fast transport, and later an important instrument of war, agriculture, and trade. Still later, this partnership would inaugurate a new concept in supportive therapy for people with special needs, limited physical abilities [2, 3], and scientific research.

The horse ancestor, *Hyracotherium* (also known as *Eohippus*), dates from the Eocene period [1] (about 55–35 million years BP=before present, for modern archaeologists) and lived in the

northern hemisphere; the genus *Equus* appeared later in North America and migrated to Asia in the Pleistocene [1] (about 2.6–0.12 million years BP). It is believed that the wild coat color (more frequently from occurring in nature; “general rule”) of the first horses was a light brown body (ranging from yellowish to light brown) with dark mane, tail, and limbs plus a dominant dilution called dun (characteristic marks and dilution of the body color); this phenotype apparently provided satisfactory camouflage against predators. At the end of the Pleistocene and early Holocene, appaloosa and black coats were already existent [4] in primitive herds.

Throughout horse’s history, appearances of different phenotypes were promoted by genetic mutations. Such different phenotypes became more frequent when equine populations faced climatic and geological events of major proportions. For example, appearance of black horses at the beginning of Holocene in Europe (eastern and central regions) and Siberia could mean that there was a postice age migration and subsequent selection due to increasing foresting of areas. Reducing the space occupied by a given population would have led to departing from Hardy-Weinberg equilibrium (principle of gene balance) [5, 6] and the chances of mating between individuals carrying mutations increased, providing the birth of homozygotes for features that were previously not expressed in the phenotypes. But it was only with domestication, and later, when horse breeding was established in primitive societies (confinement of an equine population in smaller spaces with restricted matings; slaughter of undesirable individuals, as male offsprings and other stallions) that an explosion of new colors began. Black and chestnut horses became more frequent in the Copper Age, and it was not until 3000 years BP, at the beginning of the Bronze Age, that today’s most widespread-diluted coats appeared [4].

It is also important to mention that at the time the horse was the main means of transport by land, different colors draw the attention of people holding some kind of power (political, financial, or religious) such that many understood that riding striking-looking animals would also represent the power those people possessed. As the Australian poet Pamela “Pam” Brown wrote:

“The horse is the projection of people’s dreams about themselves—strong, powerful, beautiful, and it has the capability of giving us escape from our mundane existence.”

Thus, kings, tribal chiefs, senior military ranked officers, noblemen, emperors, and the great conquerors were generally seen mounting horses of exquisite colors, which distinguished them from the common people. Those powerful ones usually rode stallions, and so it becomes easy to understand the spread of these mutations in relatively reduced timespan: a mare produces, even considering the advances in reproductive biotechnologies of today, much less offspring per year than a single stallion can produce by the same time.

Nowadays, various breeds are known to have impressive looking animals (although aiming a performer), either through changes in the hair structure (such as the magnificent metallic sheen of the Akhal-Teke or the curly hairs of the Curly breed), or through different meshes and patches scattered over the horse’s body. Usually, the colors with greater visual appealing are those originating from the combination of several mutations that lead to the localized

absence of pigment producing cells, contrasting with areas of intense and darker colors. Some of these mutations deliver somewhat unpredictable effects but still have great visual appeal.

## 2. Basic concepts

Since the dawn of horse domestication, coat colors have always fascinated mankind. However, it was not before a century ago that the knowledge of genetics and the development of scientific tools have become powerful enough so that major changes could be identified and their effects better understood.

Much of the confusion arising when the purpose is naming a coat color refers to the fact that not always this classification is in accordance with the genetic origin of the color or colors in question. For example, a very dark stallion mistakenly registered as dark brown, could be, in fact, a very dark chestnut. Another extreme example would be the presumed occurrence of a third allele at the Extension locus (other candidate locus might be the beta defensin locus [7]), relatively rare, called dominant black (represented as  $E^D$ ) [8]; considering expression of only Extension and Agouti loci, such allele, when present, would determine an entirely black phenotype, regardless of the existence of the dominant allele agouti ( $A^A$ ) on the ASIP locus (which encodes a bay phenotype in eumelanistic horses). The important information about dominant black here is, no matter the intensity of the black (some could display even a near-black brown) color on a dominant black horse, it completely overrules the ASIP locus expression.

Another example would be the genetic difference between roaned (sparse white hairs intermingled with the basic color) and the true roan (sparse white hairs intermingled with the basic color, while maintaining solid mane, tail, limbs, and head): While no specific locus (or loci) has yet been determined for the roaned trait—in fact, it may even be not an inheritable characteristic [9]—the true or genetic roan is clearly documented as a dominant mutation in the KIT ( $KIT^R$ ) [10] gene.

Additional difficulties on identification arise when considering:

Different classifications according to each breed standard. Different nomenclatures for the same coat color are not rare, depending on different countries, or regions within the same country. For example *alazão sopa de leite* (Southeastern Brazil), *baio ruano* (Southern Brazil, Uruguay, Argentina, and other Spanish-speaking countries), and *palomino* (USA and other English-speaking countries) are different names standing for the same genotype, which in time codes for a yellowish coat color with lighter mane and tail.

Subjective perception of a particular shade—for example, mahogany bay, dark bay, light bay, blood chestnut, liver chestnut, chocolate palomino, and jet black.

In fact, for some colors the expression of many genes plus external influences and individual factors (such as animal's age, time of year, insolation, and nutritional status) can make it somewhat difficult to identify a horse's color. A classic example is a black horse kept outdoors

all summer long, which causes a “fading” effect on coat color, making it easily confounded with a dark brown horse to the eyes of an unaware observer.

To really understand the genetic basis of coat color in horses, it is necessary to first understand some basic precepts. Because of empirical observations and secular traditions, it is not always easy to accept some of these key concepts:

Even though a wide range of nuances are noticeable in coat colors, and even considering the different classifications according to racial standards and individual assessments, the melanocytes of the horses are able to produce only two types of pigments [11]: eumelanin (black pigment) and pheomelanin (reddish pigment). Melanocytes are cells with dendritic extensions, originated from the embryonic neural crest and in which melanogenesis takes place in specialized organelles called melanosomes [12]. From these two pigments and the influence of several other genes [13], all variations of colors observed in the horse coat color are produced. For melanogenesis to happen, specific enzymes are required in melanocytes, being tyrosinase the one responsible for the initial step in the melanogenesis process, converting the amino acid tyrosine into a compound called dopaquinone [11].

The important matter herein is the definition of “points” of a horse. These mean manes, tail, and lower part of limbs and to some scholars, the inner line of the ear. For the identification of coat color it is very important to understand the concept of “points,” as this is the first step to identify the color genetics: horses may have black points or points of the same color of the rest of the body (in some cases there can be variation on shade, though the color is the same), in the case of a nonblack horse. By points of the same body color we mean they are usually yellowish or reddish, but may also be dark brown (if the horse is brown). Knowing the horse’s points color will provide the differentiation between horses producing eumelanin (black, dark pigment) from those producing only pheomelanin (reddish pigment). That is, horses with black points, independent of the rest of the body color, have at least one dominant allele at the Extension locus:  $E^E$ . Animals with points of the same color as the rest of the body (except in cases of entirely white horses) have a double dose (homozygous) of the recessive allele at the Extension locus:  $E^eE^e$ , meaning it is impossible to them to produce the black pigment eumelanin [14, 15].

The white in the coat of a horse is always superimposed on the basic coat color. The horses are primarily colored, being capable of producing pigment all over the body, and any and every white in the equine species is derived from a genetic directive that prevents color development due to the absence of melanocytes in that part of the body [16], and not to the lack of pigment production by melanocytes. Whether it is the meshes of a tobiano, or the entirely white body of an extensive expression of a white spotted, this means that in these areas there are no melanocytes. The same goes for the common white marking seen on the face and lower portion of horse’s limbs (e.g., markings commonly found on the head as “star,” “stripe,” and on the limbs as “socks”). Such common marking is also due to the absence of melanocytes in the white hairs regions and has polygenic origin influenced by other random factors [13].

Once understood these fundamental concepts, one may proceed to the study of the genetic basis of the several coat colors.

### 3. The three basic colors

Three basic colors, built from the two pigments produced by horses, are derived from the interaction of two genes: melanocortin-1 receptor (MC1R) and agouti signaling protein (ASIP) [14]. These colors are *bay*, *black*, and *chestnut*. The influence of other genes (for example, dilution genes) determines the variations of each of the three basic colors [13].

-MC1R gene: The dominant allele ( $E^E$ ) determines the production of eumelanin; the recessive allele ( $E^e$ ) determines the production of pheomelanin. The MC1R gene encodes the production of a protein called melanocortin-1 receptor, which plays a pivotal role in the pigment production process. This receptor is located on the surface of melanocytes. When activated by melanocyte stimulating hormone (MSH), it triggers a series of chemical reactions within the melanocytes, leading to production of eumelanin. In fact, melanocytes are capable of producing both eumelanin and pheomelanin, but when the animal has the  $E^eE^e$  genotype, what happens is that the receptor is defective, making it unable to properly transmit the information passed by MSH, thus leading to production of pheomelanin only. The net result of the forms of this gene is therefore an animal with a black coat when there is at least one dominant allele ( $E^E E^E$  or  $E^E E^e$ ), or a *chestnut* horse (reddish color) when homozygous to the recessive allele ( $E^e E^e$ ). The recessive allele ( $E^e$ ) is a mutation of the wild form ( $E^E$ ), which occurred about 7000 years BP [1].

-ASIP gene: the dominant allele ( $A^A$ ) encodes production of a protein called agouti signaling protein—hence the gene's name—which has the ability to block the melanocortin receptor 1 in existing melanocytes in the body (but not in the points) of the horse. When the receptor is blocked, there is no stimulus for the production of eumelanin (as explained above), and only pheomelanin synthesis occurs. Therefore, an eumelanic horse ( $E^E E^e$ ) with a dominant allele at the Agouti locus ( $A^A A^a$ ) yields a *bay* horse—a horse with a reddish to brownish body color and black points. About 8000 years BP [1], a recessive mutation occurred in the ASIP locus, yielding black color throughout the entire body of the horse. For this to happen; however, two doses (homozygosity) of the recessive allele are required:  $A^a A^a$ . An interesting fact is that the agouti signaling protein is effective only in eumelanin-producers horses (i.e., whose genotype is  $E^E E^E$  or  $E^E E^e$ ). As in *chestnuts* (whose genotype is  $E^e E^e$ ) melanocortin-1 receptor is defective, the protein has no effect.

As a net result of the interaction of MC1R and ASIP genes, we have:

*Genotype  $E^E E^- A^a A^a$  = black horse.* It is a very common color in the Percheron breed, and valued when having white markings in some other breeds. Some breeds, such as Friesian admit only black animals. Note that a superscript dash character (-) along with the locus letter (as in  $E^E E^-$ ) does not mean “negative,” this dash means that the allele occupying that locus is unimportant to the phenotype.

*Genotype  $E^E E^- A^A A^-$  = bay horse,* comprising all its variants regarding shades: *Light bay*, *dark bay*, *blood bay*, *mahogany bay*, etc.; in South American countries, a bright red bay is called *colorado*; a faded shade of *colorado* is called *douradilho*. There are two more alleles (so far) possible to exist at this locus:  $A^+$  for *wild bay* (a bay where black on the limbs is restricted to the region below the fetlocks) and  $A^t$ , which would be responsible for the *near black* coat

color with conspicuous lighter regions. These lighter areas are tan or reddish, and should not be confused with the mealy modifier ( $Pa^+$ ), a dominant trait that affects all the basic coats, causing paler areas around the eyes, muzzle, distal part of limbs, and underside the body. In Latin America countries, this variant is known as *zaino* (causing many mistakes by the time of breed registry); in English-speaking countries they are called *seal browns*.

*Genotypes  $E^eE^e A^A A^-$  or  $E^eE^e A^a A^a$  = chestnut horse.* This means that *bay* is, in fact, hypostatic to *chestnut*, that is, the dominant allele in the locus Agouti ( $A^A$ ) is masked by the homocigosis of the recessive allele in the Extension locus ( $E^eE^e$ ). In other words,  $E^eE^e$  results in chestnut horses regardless that alleles are present in the ASIP locus.

Although recessive homozygous alleles at the MC1R locus encode always pheomelanin animals, each country (and regions within the same country) has its own designation based solely on phenotypic shade of red in the coat. For example, in Spanish and Portuguese-speaking countries, there is the *alazán*, *tostado*, and *ruano*, each one becoming even more detailed when more variants are perceived on the basic shade. Examples: *Tostado negro*, *tostado requemado*, *alazán sangre*, and *alazán tostado*. In some English-speaking countries, the term *chestnut* defines a darker red horse, while *sorrel* would be a lighter colored red horse. Some breed associations, such as the American Quarter Horse Association (AQHA) use this distinction as a rule. Some breeds, namely the Suffolk Punch and the Haflinger admit only chestnut or sorrel animals.

As for breeders of draft horses, *sorrel* would be applied to a chestnut horse carrying the mealy modifier.

Once determined which pigment(s) is (are) produced by the horse and understood the relationship between all possible alleles in genes MC1R and ASIP, it is possible to understand the effects that other alleles, at other loci, yield over the three basic colors.

## 4. Diluted colors

Occur when proteins encoded by specific genes alter amount of pigment or deposition of pigment in receiver structures (keratinocytes and hairs). The effect is a dilution in the intensity of the original color.

### 4.1. Dun

(*gateado*, in Latin American countries): This is perhaps the oldest dilution, better known by the so-called *primitive marks*. It is believed that it was present in the primitive herds, being an effective camouflage against predators (**Figures 1 and 2**). Thus, the genotype notation would be the wild-type  $D^+ D^-$ . It is a consistent coat color in the Tarpan (*Equus ferus ferus*), Przewalski (*Equus przewalski*), and breeds like Konik (only *grullos*, black animals with the dun dilution), Sorraia, Fjord, and Dülmen.

The first mutation encoding a nondun phenotype occurred about 45,000 years BP, that is, predating domestication of horse. This phenotype retains the primitive marks, though a little lighter than the wild type. Evidences suggest a second nondun mutation originating from a chromosome



Figure 1. Dun (gateado, in Latin American countries): the oldest dilution factor.

carrying the first nondun mutation and that would have happened only a few thousand years ago; this is consistent with selection to colors with no camouflage at the time of domestication [17]. The dun effect over the basic colors—no matter how many dominant alleles are present—is a dilution of body color but not of head and points color, and the primitive marks, being the *dorsal stripe* the more consistent and best observed one. Primitive marks vary in extent and intensity. They are called so because are displayed in breeds considered primitive, though not limited to. They consist in darkening of coat on some body parts, forming recognizable patterns:

- *Cob webbing*: These are thin darker stripes on the horse's front. It has radial orientation, hence the name *cob webbing* or *spider webbing*. It is not visible in all duns, being more visible in *grullos* (black animal carrying the dun dilution).
- *Shoulder stripe*: A shadow or stripe of darker hairs running distally from (and over) the withers. It may vary in color intensity, shape, and extent.
- *Zebra stripes, stripped legs or leg bars*: Horizontal stripes appearing on the carpi and hocks, varying in number and color intensity. These are not always seen. In *grullos* they are usually black or very dark colored.
- *Dorsal stripe*: This is the most common mark on the dun dilution carriers. If a horse lacks the dorsal stripe, then it does not have the dun factor. It is a dark stripe clearly distinct from the rest of the coat, extending from the poll to the dock of the tail on the



Figure 2. Dun markings as zebra and shoulder stripes are shown in more detail. *Source 1 and 2: Mr. Jose Victor Isola.*

dorsal line. Some horses not carrying the dun dilution may display a dorsal stripe, though this is less clear and lighter, and often subject of controversy. Whether this type of dorsal stripe is due to the first of mutations in wild-type allele [17], remains to be cleared by future researches.

- *Facial mask:* It is the undiluted color on the head coat. It varies in size, from a small band over the bridge of the nose or forehead to an entire undiluted head coat.
- *Guard hairs* (sometimes called *frosting*) are lighter colored hairs growing on the edges of the mane and at the tail base. A good example is the Fjord breed. Despite having different characteristics, care should be taken to not confuse guard hairs with those found on the tails of *roans* (genetic roans or not) and *rabicanos*.

There are some primitive marks that are not observed in all horses carrying dun, or slightly observed, or observed only during the first months of life. Such marks are not characteristic nor they attest the presence of the dun factor by themselves; often they are not even present in all duns. These appear as:

- *Dark bands* perpendicular to the dorsal stripe (barbs off the dorsal stripe) that may extend variably toward the ribs;
- *Shading on the neck (neck shadow)*, which can vary from a well-defined mark to only a shadowing;

- Small dark marks (roughly the size of a pea) called *mottling*, on the stifle, legs, shoulders, and arms;
- *Dark bars on the ears*, occurring below a dark tip of the ears (upper third of the ear, caudal view), which are observed in most duns;
- *Zippers*, a thin longitudinal line of lighter hairs along the shins;
- *Bider marks*—a dark band or shadowing varying in shape and extent on the shoulder, identified so far only in some horses from Mongolia [18].

The nomenclature question when it comes to dun dilution colors, is somewhat complex due to different countries and breed registries where it is permitted. In Latin America countries these animals are called *gateados* preceded or succeeded by a term that identifies the underlying coat. In North America they are called *linebacked duns*, a clear reference to the typical dorsal stripe of the dun dilution. A wide range of terms is assigned to the various nuances caused by dun on basic coats and other dilutions. Interestingly, often two or more nuances, despite having different names, actually have the same genetic signature.

Until a few years ago, the gene responsible for the phenotype dun was unknown; it was known only about its autosomal mode of inheritance and complete dominance. Thus, it has been represented by the capital letter D, and the possible alleles as + (wild) and nd (mutant allele, not dun); with the recent release of the locus responsible for dun dilution, and possibly new genetic tests that will determine the type of mutation (nd1 or nd2) [17], this notation might be amended.

The net effects of dun on the basic coats are as follows:

- *Genotype E<sup>E</sup>E<sup>-</sup> A<sup>a</sup>A<sup>a</sup> D<sup>+</sup>D<sup>-</sup>* That means a *black* horse with dun dilution becomes a *grullo* (female: *grulla*). *Grullos* are animals of yellow-grayish coat and black points and head. Like all carriers of this dilution, the primitive markings are present. In Latin countries they are called *lobuno*, a clear allusion to the wolf color (“*lobo*” is the Spanish and Portuguese word for wolf). This is one of the most susceptible to seasonal variations phenotypes: During winter months, when the coat is denser with longer hairs, the body presents an almost uniform and lighter cream color compared to summer months when the color can come to a grayish color similar to a mouse color. Obviously, cream and gray tones vary individually according to other genes in action, as well as the factors already mentioned, such as nutritional status and age. Thus, the lighter shades of this variant in North America are designated *silvery grullo*, *olive dun*, or *slate grullo*; the darker shades are called *wolf dun*.
- *Genotype E<sup>E</sup>E<sup>-</sup> A<sup>A</sup>A<sup>-</sup> D<sup>+</sup>D<sup>-</sup>* This is a *bay* with dun dilution: In Latin America countries, it is called *gateado*; in North America nomenclature varies greatly, richly describing the intensity of the underlying color. Thus we have the *golden duns* and *silvery duns* (dun on *light bay*), *zebra dun* (dun on *medium bay*), *dark dun* (dun on *dark bay*), and *coyote dun* (dun on *sooty bay*); *dusty dun* (dun on lighter shades of *seal brown*), *shaded mealy grullo* (dun on intermediate shades of *seal brown*—note that *mealy* and *grullo* here are solely

phenotypic definitions, that hold no genetic match with the real (genetic) *mealy* and *grullo*, *shaded wolf dun* (dun on darker shades or sooty variant of *seal brown*). Sooty is the presence of very dark hairs spread over the coat, especially on the dorsal region. Many theories offer explanations for its origin and genetic mechanisms, but none has yet been adequately proven.

- *Genotype E<sup>e</sup>E<sup>e</sup> A<sup>+</sup>A<sup>-</sup> D<sup>+</sup>D<sup>-</sup> Chestnut* carrying dun: In Latin America countries it is called *gateado ruivo* (something like “auburn dun” or “ginger dun”). In North America, the dun on lighter *chestnuts* (and *sorrels*) is called *apricot dun* or *sooty apricot dun* (for sooty variant of *sorrel*), while dun on *tostado* (*chestnut*) ranges from *apricot dun* or *claybank dun* (lighter shades) to *red dun* for intermediary and sooty shades of *chestnut*.

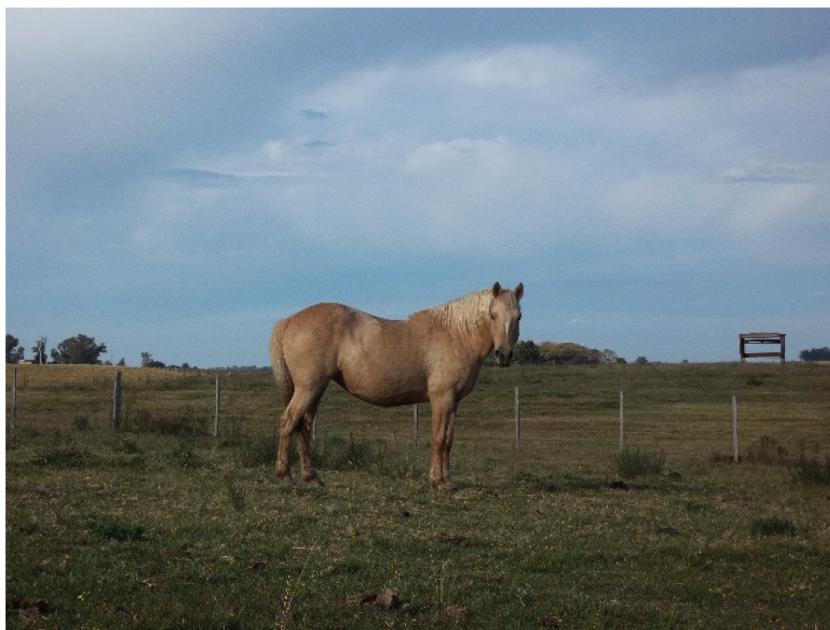
#### 4.2. Cream

Another widespread dilution; results from a dominant mutation in the membrane-associated transporter protein gene (MATP) [19]. The effect over the coat color depends on the allelic zygosity—what is called *incomplete dominance*. It means that homozygosity or heterozygosity of the dominant allele affect differently the phenotype. In the case of cream dilution, it happens that homozygosity of the dominant allele affects the phenotype more aggressively than heterozygosity. The MATP gene is represented by the C letter by convention; this comes from a time when it was thought that a “C” gene (with “C” standing for “color”) was responsible for the color over the body, so that double-diluted cream horses lacked the “color gene” and were regarded as albinos. Obviously, the cream dilution had not yet been postulated at that time. Additionally, there are different effects of the cream dilution over feomelanin and eumelanin horses. Thus, the genotype C<sup>Cr</sup>C acting over pheomelanin will produce cream-colored hairs, varying from a very light cream color to a dark caramel, while genotype C<sup>Cr</sup>C<sup>Cr</sup> will dilute hairs color to an off-white shade and will dilute also the skin color (which will become pinkish); the eyes will have a bluish color. On the other hand, the genotype C<sup>Cr</sup>C will have little (if much) effect over eumelanin (*black* horses with one copy of the mutated allele may display a faded black color or, in some cases, a very dark brown color), while homozygosity will produce a horse with off-white-colored hairs, pinkish skin, and bluish eyes. Therefore, we have the following net effect of cream dilution on the basic colors:

- One gene cream on *chestnut* = *Palomino*, a valued coat color especially when the body holds a golden sheen. The body is yellowish ranging from pale yellow/cream to dark caramel. Palominos can vary in shade, though manes and tails are usually lighter like the horse in the picture. The genetic representation is as follows:

E<sup>e</sup>E<sup>e</sup> A<sup>+</sup>A<sup>-</sup> C<sup>Cr</sup>C Note that when the Extension locus presents both recessive alleles, any combination of alleles at the Agouti locus (A) will not produce any effect on the phenotype: *Chestnut* horses can only produce pheomelanin.

- Two genes cream on *chestnut* = *Cremello* (*cremello* is the definition given to the horse with very pale cream coat color, pink skin, and blue eyes). These animals usually have



**Figure 3.** Palominos may vary in shade from a very light cream color like the horse in the picture. *Source:* Mrs. Albrecht.

eye and skin problems when not protected from continuous sun exposure (**Figure 3**). These horses are called *salgo* (because of the light-colored eyes) or *melado* in Spanish and Portuguese-speaking countries. The genetic representation is:

$E^e E^e A^a A^a C^{Cr} C^{Cr}$  Again, any combinations of alleles in the Agouti locus do not produce any effect on the phenotype, because  $E^e E^e$  horses are capable of producing only pheomelanin.

- One gene cream on *black* = *Smoky Black* (the coat color is usually faded black). In some Brazilian regions it is called *preto macaco*. The genetic representation is:

$E^E E^E A^a A^a C^{Cr} C^{Cr}$  Note that in the Extension locus, only one of the alleles must be dominant to produce eumelanin. For the horse to display a black coat over the entire body (i.e., with melanocortin-1 receptors present and responsive to MSH in all melanocytes), it is necessary homozygosity of the mutant allele at the Agouti locus.

- Two genes cream on *black* = *Smoky Cream*, a very pale cream coat, with pinkish skin, and blue eyes; however, usually some areas still retain dark pigment—seen like dark hairs scattered over the entire body. The genetic representation is:

$E^E E^E A^a A^a C^{Cr} C^{Cr}$  Note that, at the Extension locus, only one dominant allele is needed to produce eumelanin.

- One gene cream on *bay* = *Buckskin* is the name of the animal with black points and yellowish body, ranging from light cream to dark caramel. In Spanish and Portuguese-speaking countries it is called *baio*. The genetic representation is:

$E^E\ A^A\ A^- C^{Cr}C$  Note that, at the Extension locus, only one dominant allele is needed to produce the agouti signaling protein.

- Two genes cream on *bay* = *Perlino* (very pale cream coat with pinkish skin and blue eyes; some areas still retain dark pigment, mainly on the points. Also called melado and salgo in Spanish/Portuguese-speaking countries. The genetic representation is:

$E^E\ A^A\ A^- C^{Cr}C^{Cr}$  In Extension and Agouti loci, it is necessary only one dominant allele to produce eumelanin and agouti signaling protein, respectively.

Cream-diluted coats are highly valued by some breed associations, for example, the American Cream Draft Horse.

Because dun and cream loci operate separately, it is possible for a horse to carry both dilutions. In this case, the colors become even more diluted and an identification based solely on the visual aspect becomes more difficult and prone to mistakes.

#### 4.3. Silver or silver dapple

It is a mutation at PMEL17 gene [20] (premelanosomal protein 17), which encodes a transmembrane protein involved in melanosomal production of eumelanin. The mutation produces a defective protein, causing a dilution effect on that pigment. Hence, the black color is diluted to shades of sepia or chocolate (called *silver chocolate* in some locations; in Australia it is known by *taffy*) and the manes, tails, eyelashes, and tactile hairs of face and muzzle show a grayish to whitish, sometimes silvered color. Limbs can be either whitish or near black colored. Usually the coat is dappled, and the more heavily dappled, the more valued it is. It is a dominant trait represented as  $Z^Z$ ; the recessive allele (wild, normal, and nonsilver) is represented by  $Z^+$ . The variants are named according to the intensity of dilution; so we have the *blue silvers*, *black silvers*, *silvers chocolate*, and *silver dapples*. In the case of bay horses carrying the silver dilution, only eumelanin regions are affected, and so they are called *red* or *bay silvers* (in Rocky Mountain Horse breed, these are called *red chocolate*). It is a highly valued coat in some countries and quite common in breeds such as the Rocky Mountain Horse, Mountain Pleasure Horse, Kentucky Mountain Saddle Horse, and the Icelandic breed. It also appears in Quarter Horses, Morgans, and Miniatures. In horses homozygous for the condition it has been observed different combinations of symptoms affecting the anterior and posterior segments of the eye; such condition is known as multiple congenital ocular abnormalities (MCOA) [21–23].

In the same way, the cream and dun dilutions can act together, so when  $Z^Z\ Z^+$  occurs on a diluted coat color (by any other diluting gene) the degree of difficulty of color identification increases, and cases of mistaken registration due to several concurrent dilutions are not rare. For example, a *silver buckskins* can be easily confused with a *Palomino* due to the golden color

of the body and pale points; *silver buckskins* usually have lighter skin, which can make them superficially similar to some *champagnes* (another dilution).

#### 4.4. Champagne

A dominant mutation in the SLC36A1 gene (Solute Carrier 36 family A1) showing complete dominance [24]. It causes dilution of coat color, sometimes giving it a metallic sheen; it also makes the skin to dilute to a light brown color (with many small and darker splotches like freckles) and amber eyes. The dilution is represented as  $Ch^{Ch}$  (mutant allele) and  $Ch^+$  stands for the wild-type allele, nonchampagne. The net effect on the basic coat colors completely changes the original names. Sometimes very clear *champagnes* are confused with some cream-related coats colors due to a very light shade presented in some variants; however, the striking characteristics of skin and eyes colors should avoid confusion. Champagne dilution causes dilution of the coat color ranging from a very light yellowish to a light chocolate shade with varying points colors. The genetic representations are:

$E^E E^- A^a A^a Ch^{Ch} Ch^+$  it is called *classic champagne*. The color varies from dark beige to light chocolate, with points slightly darker than body.

$E^E E^- A^A A^- Ch^{Ch} Ch^+$  Champagne on a *bay* coat. It is called *amber champagne*. The body color is light tan brown or light yellowish brown and the points are usually light chocolate colored. Champagne on a *seal brown* coat ( $E^e E^e A^t A^{t/a} Ch^{Ch} Ch^+$ ) or on darker shades of *bay* is called *sable champagne*.

$E^e E^e A^- A^- Ch^{Ch} Ch^+$  This is a *gold champagne*, often confused with *palominos* due to the golden glister and lighter color. The combination of champagne with a cream allele causes an even lighter color, with greenish blue eyes with a hint of amber—these horses are called *ivory champagnes*.

#### 4.5. Other dilutions

There are other dilutions found hitherto in some breeds (some not yet fully characterized from a genetic point of view). Examples are the *pearls* (dilution similar to champagne), *mushrooms* (dilution similar to silver, though it affects pheomelanin instead of eumelanin; so far found only in Miniatures). Some of these alleged dilutions/mutations have not yet been critically studied on a scientific basis. A dilution known as the lavender foal syndrome occurs in the Arabian breed and is due to a defective Myosin Va protein, coded by a recessive mutation in the gene MYO5A [25], causing severe neurological disorders and very light reddish, almost lilac hairs and skin in the neonate. These and other dilutions are an interesting and extensive subject, which could easily be the subject of an entire chapter.

### 5. Coat color modifiers

These genes impose areas of different colors on the basic coat color, modifying the original aspect.

### 5.1. Pangaré or mealy

This is a modifier whose genetic origin is unclear. It makes some areas paler than the rest of the body. These are around the eyes, muzzle, underside the body, flanks, and inner part of limbs [13]. The name pangaré derives from the observation made in horses of Spanish-speaking countries, and is translated into English as mealy. Apparently it is a dominant trait and is considered a wild trait, i.e., likely to be present in the original color of primitive horses. It is represented by the notation  $Pa^+$ , with the recessive mutation represented by  $Pa^{np}$  (standing for nonpangaré). It affects all coat colors, with its most dramatic effect seen on the *sorrel* coat color. Generally mealy *chestnuts* or *sorrels* (in Spanish and Portuguese, only *pangaré*) present a mane very lighter than the rest of the coat. The body color becomes slightly lighter, although pangaré is not considered a dilution factor. In *black* horses, it renders the coat to a slightly lighter or brownish color, easily leading to confusion with the *seal brown* color; it differs from the *seal brown* primarily due to the shade of the light areas: while in *seal brown* the lighter areas are tan and less extensive, in the *pangaré* those areas are more extensive and pale.

### 5.2. Rabicano

Its mechanism of action and genetic control are still unknown [13]. It causes the appearance of white hairs scattered on the coat over the rib cage (indeed following the direction of the bones), flanks, and tail base. It varies in intensity of expression and may appear as only a few white hairs on the base of the tail and sides of the rib cage, to a large amount of white hairs over these regions, in this case being easily confused with roaned or a true roan.

### 5.3. Flaxen

It is a recessive trait with variable expression; when homozygous determines a lighter shade (than the body) of the manes and tails (Figure 4). Acts only on chestnuts and sorrels. Its mechanism of action and genetic control are unclear, and it is even accepted to be of polygenic nature [13].

### 5.4. Brindle

Brindle is a very rare color, referred to as vertical dark stripes on any background color [13]. Usually, points and head are unaffected and may occur in any breed. It was suggested two causes for the brindle in horses [13]: either it is a reorganization of sootiness into vertical stripes instead of the usual sooty variant, or it is a chimera; in fact, some animals have been proven to be chimeras, resulting from the fusion of two embryos early during pregnancy.

## 6. The gray case

Gray is a dominant mutation in the gene STX17 [26], Syntaxin 17, represented by the capitalized letter G. It consists in a gradual and progressive whitening of all hairs on the horse's body, but the skin may remain dark for many years. In some cases, such as the Lipizzaners, the skin becomes gradually unpigmented as well. When this mutation is present, at every shedding time

the original color of the horse will be replaced by white-colored hairs, even if the underlying coat is determined by dominant alleles. Thus, the gene that causes the graying of the coat is epistatic, which occurs when a gene at one locus masks the effect of another gene at another locus.

The dominant nature of the mutation implies that any gray horse is required to have at least one of its parents being a gray too. Gray horses, especially homozygotes ( $G^G G^G$ ) tend to develop a type of tumor called the melanoma of the gray horse [27].

Animals are born colored (the color of the original coat) and graying starts around the eyes, already during the first months of life. However, total whitening can take years, ranging individually from horse to horse. This mutation leads to an excess of melanin production, and that is the reason why gray-carrying horses are born with vivid darker colors; such a characteristic is well observed in the foals whose coat is a jet black, unlike the black foals without the mutation, who are born a grayish color. This is the reason because some breeders say "blacks are born grays and grays are born blacks."

Many denominations are applied according to different stages of the graying process, varying by country: *Tordilho medalhado* or *apatacado* (*tordilho/tordillo* stands for gray in Portuguese and Spanish; respectively, the word is derived from "tordo," a bird whose plumage has a similar color) is the same of *dapple gray*; *branco/blanco porcelana* is the Portuguese/Spanish equivalent of *white porcelain/porcelain gray* (older grays with depigmented skin contrasting with the white hairs of body), *tordilho/tordillo vinagre* (*rose gray*), *tordilho/tordillo negro* (*iron gray*), and *tordilho/tordillo*



Figure 4. A cremello horse has pinkish skin, blue eyes, and a coat of a very light shade of cream. Source: Mrs. Albrecht.

*pedrês (fleabitten gray)*. It is a widespread color between various breeds such as Thoroughbred, Purebred Arabian, Percheron, Andalusian, and it is the only color seen in Lipizzaners.

## 7. White patterns

Consist of white hairs mixed by or arranged in white spots overlapping the underlying coat; it has dominant or polygenic nature and arises from mutations in several genes. The phenotypic appearance can vary from one extreme to another, often being impossible to say how many and which mutant genes are present without genetic testing. These white patterns should not be confused with common white marks present on the face and the lower portion of limbs (white markings of small size, limiting white to small patches in the forehead, and white lower limbs not reaching carpi and hocks).

### 7.1. Overo

Overo frame is the name of a mutation in the gene EDNRB [28] (endothelin receptor B), represented by the letter O with superscript  $lw$  (lethal white). Much confusion has been caused due to this gene. Because it does not manifest phenotypically in all horses carrying the mutation, it was mistakenly understood as a recessive trait. For this reason, it is known as the “ninja gene” or “007 gene.” What happens though, is that  $O^{lw}$  has highly variable expression, depending on the action of other genes, which makes confusing its mechanism of action—It can even crop out in a foal from parents with apparently solid-colored coats—in these case, hiding the gene (“ninja gene!”). What is known for sure is that it is a dominant mutation capable of amplifying the expression of other white patterns such as Splashed White and Sabino.

To breeders and horse enthusiasts, *overo* is the general empirical name given to animals showing many asymmetrical white spots of relatively reduced size and jagged edges distributed throughout the body (Figure 5); however, only with genetic testing one can be sure that the frame gene is present.

Usually limbs are solid (i.e., do not have any white marking) and the head is largely marked with white (sometimes called “baldface”). Irregular white spots appear on the sides of the body and neck (whence the name “frame”); many *overos* has the upper lip solid colored and blue eyes are common, even when the coat around the eyes is pigmented. *Overos* minimally marked can show only the baldface.

The problem with overo frame is that it is lethal when homozygous, causing the lethal white foal syndrome [28]. The medical condition is called ileocolonic aganglionosis and requires euthanasia of the foal by the first days of life for humanitarian reasons. It is a pleiotropic effect of the mutant gene EDNRB, resulting in faulty areas of intestinal innervation and a foal completely (or near so) white colored. Pleiotropy is a genetic condition in which a gene causes more than one effect [29]. That is why breeding programs discourage mating of two *overos*. Making use of the Punnet square, a tool that allows simple calculation of outcome probabilities out of a given mating, the possibility of obtaining a homozygous overo frame ( $O^{lw} O^{lw}$ )—a



Figure 5. Flaxen. Source: Mrs. Albrecht.

lethal white foal—from the mating of two *overos* is 25% which is an unacceptable risk both for ethical and economical issues.

## 7.2. Splashed White

At least two loci are identified as having mutant alleles for Splashed White [30]: microphthalmia transcription factor (MITF) and paired box-3 (PAX3). Three mutations have been identified hitherto, SW1 (dominant mutation in MITF gene), SW2, and SW3 (dominant mutations in the PAX3 gene). Splashed White causes white spots in a horizontal pattern on the horse's body coat, and these white spots always cross the ventral line. The head is extensively marked (the white mark on the forehead and muzzle is quite wide and round and the muzzle is usually all white; this is called a “blaze”) and the limbs are white above carpi and hocks. In extreme-marked horses, the head may be entirely white and the body may keep color only on the dorsal line. The more copies of mutant genes, the greater the amount of white. Splashed Whites minimally marked can display only an extensively marked forehead, nose, and muzzle; limbs can have white only below shins, and usually in the forelimbs.

In Latin America countries, it is called *bragado* preceded by the name of the main coat (i.e., *colorado bragado*). In English-speaking countries, it is used *splash* after the name of the main coat, as in *bay splash*.

Splashed White is widespread among several breeds, including Welsh Pony, Icelandic, Finnish Draft, and Paints. It is present in some other breeds, although the respective stud books do not recognize and does not register it as a specific characteristic.

Accordingly to the extent of depigmentation on the animal's head, it may have blue eyes and/or congenital deafness. A problem of embryonic death occurs when both PAX3 alleles are mutated.

### 7.3. Roan

It is characterized by white and colored hairs intermingled over the coat in different proportions only on the horse's body, not affecting points and head. Sometimes when there is very little amount of white, it is confused with *roaned*, which does not have the same genetic origin of the true or genetic *roan*. Some confusion may also occur between *roans* and *rabicanos*. Roan is a dominant mutation in the KIT gene [10] (mast/stem cell growth factor receptor; v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog), represented by KIT<sup>R</sup>. Contrary to what was once believed, homozygosity is possible in living animals and does not cause embryonic death. It can occur on any basic coat colors, diluted or modified. The nomenclature varies according to the country of origin: in Latin America countries, the word *rosilho* (*rosillo* in Spanish) is used with the name of the main coat, as in *rosilho gateado* (a roan dun horse), *rosilho prateado* (*chestnut* with large amount of white hairs intermingled with the basic color; does not mean the presence of the silver dilution); *mouro* is the roan horse whose basic color is black. In English-speaking countries the word roan is used after the main color:



Figure 6. Overo. Source: Mr. Isola.

*red roan* (a *chestnut* or *sorrel roan*), *blue roan* (the *roan black*), *bay roan*, *strawberry roan*, and so on (Figure 6).

#### 7.4. Tobiano

Tobiano spotting is another mutation in the KIT gene [31], characterized by white spots in a vertical pattern crossing the dorsal line (unlike Splashed White, wherein the spots have a horizontal pattern and cross the ventral line). It has dominant nature and is represented by  $KIT^{To}$ . Usually the head is solid or has minimal white marking, the limbs have high socks (white mark that reaches the carpi and hocks), and the tail and manes are usually bicolored. The maximally marked *tobiano* may have the body almost completely white (especially if Splashed White is present), but keep solid-colored ears, what is called “Medicine Hat.” On the other hand, minimally marked *tobianos* may display only high socks. In these cases, only a genetic test can demonstrate the presence of this mutant gene.

A frequent feature in homozygous *tobianos* ( $KIT^{To}$   $KIT^{To}$ ) is the so-called “paw prints,” “cat tracks,” “bear paws” (English), “marcas do perigueiro” (Portuguese): colored small rounded spots arranged in small groups spread over white areas on the horse’s body. Despite the presence of these marks be of great help when one wants a *tobiano* foal, the absence of these marks does not mean the *tobiano* parent is heterozygous.

Tobiano derives its name from Tobias, a Brazilian general who came to Argentina in the mid 1800s bringing with him soldiers mounted in very odd-spotted horses. The nomenclature varies according to country. Moreover, in Latin countries it varies also among countries’ regions. *Tobiano* is used, in North and South America, including Southern Brazil, along with the main coat name: *tobiano colorado* (Portuguese and Spanish), *bay tobiano*, and *black tobiano* (English); in other regions of Brazil, *pampa de baio* (or other coat color) is used to designate a *buckskin tobiano* animal with large proportion of white spotting, or *baio pampa* (or other coat color) for the animal with little white spotting.

#### 7.5. White spotted

Formerly known as Dominant White, it is another dominant mutation in the KIT gene [16] that holds extensive phenotypic variation. It is characterized by large amounts of sparse white hairs with no set pattern, overlapping the adjacent coat. The horses also have depigmented skin and dark eyes, and in extreme cases the animal can be completely white being easily confused with an older gray. When the eyes have colors other than dark brown, there is another gene (or genes) acting. White spotted is represented by the capitalized letter W followed by a number, which corresponds to the mutation variant [32] (for example,  $KIT^{W1}$  and  $KIT^{W19}$ ). There are by now 20 known mutations, some of which are derived from spontaneous mutations in the gene. It occurs even in the Thoroughbred breed, and some countries such as Australia, New Zealand, and Japan recognize and do list these horses as authentic Thoroughbreds.

It has long been discussed the issue of lethal homozygosity ( $KIT^{W-}$   $KIT^{W-}$ ), taken as true for a long time. This is a debatable question considering that there are living animals whose genetic tests confirmed the presence of two alleles W [9, 33].

## 7.6. Sabino

Sabino is a common term originated from Spanish-speaking countries, to designate a coat speckled with white hairs that do not form a specific pattern [13]. This definition is based only on the phenotype and is generally used to define a white patterned coat that not necessarily has the genetic basis of the same name gene.

*Sabino* (genetic) occurs in two forms: a dominant mutation in the KIT gene named Sabino-1 (KIT<sup>Sb1</sup> KIT) [34]. This variant, when homozygous, causes much more white on the body (incomplete dominance). The second form has a polygenic mode of inheritance [13] and results in horses completely white or near so. The eyes may be blue or dark brown. A typical sabino feature is the roaned edges of the white spots. White spots with these characteristics are common in the belly and intermandibular space. The appearance of *sabinos* may vary greatly from only common white marks on the limbs with roaned edges, to a whole white animal with very little pigmentation, whence it is called “the everything gene.” This renders identification of the underlying coat color, whether or not there is expression of other white patterns, practically impossible without genetic testing or at least, extensive knowledge of the genealogy of the animal in question.

Usually the coat name is followed or preceded by the word “sabino:” *Colorada sabina* (Portuguese and Spanish); *bay sabino* (English).

An interesting fact about KIT mutations is that only two mutations can be presented in the KIT gene in a particular animal, since there are only two copies of the gene in every horse; that being understood, it becomes easy to see why we will never have one “white-spotted red roan tobiano” (it would require three KIT gene copies) or a homozygous tobiano roan.

## 7.7. Appaloosa

Characteristic coat color of the homonymous breed and frequent color in breeds like Knabstrupp, Miniatures, Noriker, and others. It is a white pattern with symmetric white patches. *Appaloosa* is the generic name of the leopard complex; “complex” because all patterns in this color are genetically related, and “leopard” because of the most distinctive feature observed: the round-colored spots on a white area, varying in extent over the animal’s body. It is caused by a mutation in the gene TRPM1 (transient receptor potential cation channel, subfamily M, member 1) [35]. It shows incomplete dominance, represented by Lp<sup>LP</sup>. The most common appearance is a variable solid-colored anterior part of the body and the remainder covered by white (which is called *white cap blanket* or *snow cap blanket*), which may or may not display spots, depending on the genetic makeup (homozygous vs. heterozygous). The animals called “leopards,” highly valued, present a white body (including the head) covered with white round spots (leopard spots). The leopard complex can be split into several patterns, and usually at least two of these patterns are expressed in the same animal. These patterns relate to quantity and arrangement of white areas on the animal’s coat:

- *Frost* describes a roaning pattern that consists of scattered white hairs over the entire body of the horse.

- *Varnish roan* is like frost, though areas over bony prominences are darker than the remainder of the coat.
- *Snowflake* occurs as small round patches of white hairs spread over the entire body. As well as frost and varnish roan, it is progressive from birth to maturity; that is why animals that have the leopard complex will present, year after year, an increasingly white-covered coat, although many horses end up stabilizing at some point in their lives, contrary to what happens in the case of gray.
- *Speckled* describes a pattern of extensive roaning with numerous small dark spots are placed on a white background.
- *Blanket* is a white covering extending variably over the horse's body (**Figure 8**); it can be as small as only a few white patches on the hips, or extensive enough to cover the entire body. When the animal is heterozygous for *Lp* ( $Lp^{Lp} Lp$ ) the blanket is covered by colored round spots that can vary in size (**Figure 7**). Homozygous ( $Lp^{Lp} Lp^{Lp}$ ) horses show no spots, no matter the extent of the white blanket. In case of an extensive white blanket, the entire body will display white, while maintaining some color at the points and head and some very small-scattered spots on some regions of the body. These horses are called *few spots* (**Figure 8**).

The extent of the white blanket is a function of other loci, known to exert modifying action on the extent of white in the body [13, 35]. One of these modifiers is named Pattern-1, a domi-



**Figure 7.** Blue roan. *Source:* Mr. Isola.



Figure 8. A white blanket may vary in extension. Source: Mrs. Albrecht.

nant trait ( $PATN^P$ ) that regulates the extent of the white blanket and acts when at least one copy of the leopard gene is present. Horses that do not have the dominant gene ( $PATN^+$ ) can be solid-colored coats without white cover, displaying only the patterns of white hairs and depigmented skin described below.

Some characteristics occur independently of other manifestations of the leopard complex, being present in all animals carrying the mutation, whether they are homozygous or not. These characteristics are shown as small spots on the skin of the nose, genitalia, anus, mouth, and eyelids, called *mottled*; when the skin is dark these spots are pink and when it is clear, the spots are dark. Also there are changes in the sclera, which is white, resembling a human eye to many people.

The limbs may display the common white marks like in any other coat color. In leopard complex carriers lacking these common marks, a blend of white and dark regions may be presented on the lower limbs. These are called *lightning strikes* or *lightning marks*, and the coronary band is generally colored. The hooves can be striped. An interesting feature of the leopard complex are sparse manes and tails, often called “rat tailed” and more strongly expressed in dark-colored hairs.

A common problem found in homozygous horses ( $Lp^{LP} Lp^{LP}$ ) is the congenital stationary night blindness (CSNB) [36].

The “appaloosa” mutation appeared about 25,000 years ago, according to drawings found in the Perch Merle caves in southwestern France. In the modern era, it was developed by Indians

of the Nez Perce tribe, from horses brought by Europeans in the sixteenth and seventeenth centuries; these tribes lived in the northwestern United States, at the banks of the Palouse River, from where the name Appaloosa is originated and consequently the name by which the leopard complex became known worldwide.

## 8. Conclusion

The colors of horses are determined by genetic factors and influenced by factors such as insulation, season, nutritional status, age, sex, and health. The naming of colors varies according to individual perceptions of shades and colors, countries (regional classifications), and breed standards. A wide variety of colors appeared after the domestication of horse, with more impressive-looking phenotypes preferred since antiquity. The cells responsible for pigment production are the melanocytes, where melanogenesis takes place in specialized organelles called melanosomes. Two pigments are produced by horses (eumelanin and pheomelanin) and the interaction of the MC1R and ASIP genes determines the three basic colors of horses (bay, chestnut, and black). These pigments can be influenced by dilution genes and modifier genes, resulting in a variety of colors. Observing the color of a solid-colored horse's points, one can say which pigment is produced. Any white in the coat of a horse derives from the absence of melanocytes in the depigmented area. The exception is the process of graying, a depletion of pigment production, and leading to the gradual appearance of white hairs instead of colored hairs. Pleiotropic effects arise from the expression of some genes related to coat colors, affecting physiological systems of the horse and may even lead to death. Many colors have not been critically related to specific genes yet. Genetic tests are available to check the presence of many alleles, and without such testing is often difficult to pinpoint exactly that genes are at work to produce the color.

## Acknowledgements

The authors hereby declare, that all the pictures of horses in this chapter were taken by Mrs. Albrecht and Mr. Isola, especially on purpose to insert here, and the animals belong to them.

## Author details

Adriana Pires Neves<sup>1\*</sup>, Eduardo Brum Schwengber<sup>1</sup>, Fabiola Freire Albrecht<sup>2</sup>, José Victor Isola<sup>1</sup> and Liana de Salles van der Linden<sup>2</sup>

\*Address all correspondence to: [adripneves@yahoo.com.br](mailto:adripneves@yahoo.com.br)

1 Pampa Federal University, Dom Pedrito, RS, Brazil

2 Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil

## References

- [1] Bowling AT, Ruvinsky A, editors. *The Genetics of the Horse*. 1st ed. London: CABI Publishing; 2000. 538 p.
- [2] Ajzenmann HF, Standeven JW, Shurtleff TL. Effect of hippotherapy on motor control, adaptative behavior, and participation in children with autism spectrum disorder: A pilot study. *American Journal of Occupational Therapy*. 2013; 67: 653–663. DOI 10.5014/ajot.2013.008383
- [3] Araujo TB, Silva NA, Costa JN, Pereira MM, Safons MP. Effect of equine-assisted therapy on the postural balance of the elderly. *Brazilian Journal of Physical Therapy*. 2011; 15: 414–419. DOI 10.1590/S1413-35552011005000027
- [4] Ludwig A, Pruvost M, Reissman M, Benecke N, Brockmann GA, Castaños P, Cieslak M, Lippold S, Llorente L, Malaspinas AS, Slatkin M, Hofreiter M. Coat color variation at the beginning of horse domestication. *Science*. 2009; 324: 485.
- [5] Hardy GH. Mendelian proportions in mixed population. *Science*. 1908; 28: 49–50.
- [6] Weinberg W. On the demonstration of heredity in man. In: Boyer SH, editor. *Papers on Human Genetics*. Englewood Cliffs: Prentice Hall; 1963.
- [7] Candille SI, Kaelin CB, Cattanach BM, Yu B, Thompson DA, Nix MA, Kerns JA, Schmutz SM, Millhauser GL, Barsh GS. A  $\beta$ -defensin mutation causes black coat color in domestic dogs. *Science*. 2007; 318: 1418–1423.
- [8] Sponenberg DP, Weise MC. Dominant black in horses. *Genetic Selection Evolution*. 1997; 29: 403–408.
- [9] Holl H, Brooks S, Bailey E. De novo mutation of KIT discovered as a result of a non-hereditary white coat colour pattern. *Animal Genetics*. 2010; 41: 196–198. DOI 10.1111/j.1365-2052.2010.02135.x
- [10] Marklund S, Moller M, Sandberg K, Andersson L. Close association between sequence polymorphism in the KIT gene and the roan coat color in horses. *Mammalian Genome*. 1999; 10: 283–288.
- [11] Cone RD. Studies on the physiological functions of the melanocortin system. *Endocrine Reviews*. 2006; 27: 736–749.
- [12] Riley PA. Melanogenesis and melanoma. *Pigment Cell Research*. 2003; 16: 548–552.
- [13] Sponenberg DP. *Equine Color Genetics*. 3rd ed. Iowa: Wiley-Blackwell; 2009.
- [14] Rieder S, Taourit S, Mariat D, Langlois B, Guérin G. Mutations on the agouti (ASIP), the extension (MC1R) and the brown (TYRP1) loci and their association to coat color phenotypes in horses (*Equus caballus*). *Mammalian Genome*. 2001; 12: 450–455.
- [15] Hearing VJ, Jiménez M. Mammalian tyrosinase: The critical regulatory control point in melanocyte pigmentation. *International Journal of Biochemistry*. 1987; 19: 1141–1147.

- [16] Haase B, Brooks SA, Schlumbaum A, Azor PJ, Bailey E, Alaeddine F, Mevissen M, Burger D, Poncet PA, Rieder S, Leeb T. Allelic heterogeneity at the equine KIT locus in dominant white (W) horses. *PLOS Genetics*. 2007; 3: 2101–2108. DOI 10.1371/journal.pgen.0030195
- [17] Imsland F, McGowan K, Rubin CJ, Henegar C, Sundström E, Berglund J, Schwochow D, Gustafson U, Imsland P, Lindblad-Toh K, Lindgren G, Mikko S, Millon L, Wade C, Schubert M, Orlando L, Penedos MC, Barsh GS, Andersson L. Regulatory mutations in TBX3 disrupt asymmetric hair pigmentation that underlies dun camouflage color in horses. *Nature Genetics*. 2016; 48: 152–160. DOI 10.1038/ng.3475
- [18] Masuda M, Tsunoda J, Nomura H, Kimura N, Altangerel G, Namkhai B, Dolj U, Yokohama M. New primitive marking (Bider) in Mongolian native horses and *Equus przewalskii*. *Journal of Equine Veterinary Science*. 2007; 18: 145–151.
- [19] Mariat D, Taourit S, Guérin G. A mutation in the MATP gene causes the cream coat color in the horse. *Genetics Selection Evolution*. 2003; 35: 119–133. DOI 10.1051/gse:2002039
- [20] Brunberg E, Andersson L, Cothran G, Sandberg K, Mikko S, Lindgren G. A missense mutation in PMEL17 is associated with the silver coat color in the horse. *BMC Genetics*. 2006; 7. DOI 10.1186/1471-2156-7-46
- [21] Ramsey DT, Ewart SL, Render JA, Cook CS, Latimer CA. Congenital ocular abnormalities of Rocky Mountain Horses. *Veterinary Ophthalmology*. 1999; 2: 47–59.
- [22] Grahn BH, Pinard C, Archer S, Bellone R, Forsyth G, Sandmeyer LS. Congenital ocular anomalies in purebred and crossbred Rocky and Kentucky Mountain Horses in Canada. *Canadian Veterinary Journal*. 2008; 49: 675–681.
- [23] Andersson LS, Wilbe M, Viluma A, Cothran G, Ekesten B, Ewart S, Lindgren G. Equine multiple congenital ocular anomalies and silver coat colour result from the pleiotropic effects of mutant PMEL. *PLOS Genetics*. 2013; 8: 1–8.
- [24] Cook D, Brooks S, Bellone R, Bailey E. Missense mutation in Exon 2 of SLC36A1 responsible for champagne dilution in horses. *PLOS Genetics*. 2008; 4: 1–9. DOI 10.1371/journal.pgen.1000195
- [25] Brooks SA, Gabreski N, Miller D, Brisbin A, Brown HE, Streeter C, Mezey J, Cook D, Antczak DF. Whole-genome SNP association in the horse: Identification of a deletion in Myosin Va responsible for lavender foal syndrome. *PloS Genetics*. 2010; 6: 1–7. DOI 10.1371/journal.pgen.1000909
- [26] Rosengren PG, Golovko A, Sundstrom E, Curik I, Lennartsson J, Seltenhammer MH, et al. A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nature Genetics*. 2008; 40: 1004–1009.
- [27] Sundstrom E, Imsland F, Mikko S, Wade C, Sigurdsson S, Pielberg GR, Golovko A, Curik I, Seltenhammer MH, Sölkner J, Lindblad-Toh K, Andersson L. Copy number expansion of the STX17 duplication in melanoma tissue from grey horses. *BMC Genomics*. 2012; 13.

- [28] Santschi E, Purdy A, Valberg S, Vrotsos P, Kaese H, Mickelson J. Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. *Mammalian Genome*. 1998; 9: 306–309.
- [29] Bellone RR. Pleiotropic effects of pigmentation genes in horses. *Animal Genetics*. 2010; 41: 100–110. DOI 10.1111/j.1365-2052.2010.02116.x
- [30] Hauswirth R, Haase B, Blatte M, Brooks SA, Burger D, Drögemüller C, Gerber V, Henke D, Janda J, Jude R, Magdesian KG, Matthews JM, Poncet PA, Svansson V, Tozaki T, Wilkinson-White L, Penedo MCT, Rieder S, Leeb T. Mutations in MITF and PAX3 cause “splashed white” and other white spotting phenotypes in horses. *PLOS Genetics*. 2012; 8: 1–9. DOI 10.1371/journal.pgen.1002653
- [31] Haase B, Jude R, Brooks SA, Leeb T. An equine chromosome 3 inversion is associated with the tobiano spotting pattern in German horse breeds. *Animal Genetics*. 2008; 39: 306–309. DOI 10.1111/j.1365-2052.2008.01715.x
- [32] Haase B, Brooks SA, Tozaki T, Burger D, Poncet PA, Rieder S, Hasegawa T, Penedo C, Leeb T. Seven novel KIT mutations in horses with white coat colour phenotypes. *Animal Genetics*. 2009; 40: 623–629. DOI 10.1111/j.1365-2052.2009.01893.x
- [33] Haase B, Obexer-Ruff G, Dolf G, Rieder S, Burger D, Poncet PA, Gerber V, Howard J, Leeb T. Haematological parameters are normal in dominant white Franches-Montagnes horses carrying a KIT mutation. *The Veterinary Journal*. 2010; 184: 315–317. DOI 10.1016/j.tvjl.2009.02.017
- [34] Brooks SA, Bailey E. Exon skipping in the KIT gene causes a Sabino spotting pattern in horses. *Mammalian Genome*. 2005; 16: 893–902. DOI 10.1007/s00335-005-2472-y
- [35] Sponenberg DP, Carr G, Simak E, Schwink K. The inheritance of the leopard complex of spotting patterns in horses. *Journal of Heredity*. 1990; 81: 323–331.
- [36] Rebecca R, Bellone RR, Brooks SA, Sandmeyer L, Murphy BA, Forsyth G, Archer S, Bailey E, Grahn B. Differential gene expression of TRPM1, the potential cause of congenital stationary night blindness and coat spotting patterns (LP) in the Appaloosa horse (*Equus caballus*). *Genetics*. 2008; 179: 1861–1870. DOI 10.1534/genetics.108.088807

## Further reading

- [37] American Paint Horse Association Guide to Coat Color Genetics. Available from: <http://apha.com>
- [38] Nicholas FW. *Introduction to Veterinary Genetics*. 3rd ed. Ames: Blackwell Publishing; 2010. 329 p.
- [39] Scott DW, Miller WH. *Equine Dermatology*. 1st ed. St. Louis: Elsevier Science; 2003, 829 p.

# Major Histocompatibility Complex-Associated Resistance to Infectious Diseases: The Case of Bovine Leukemia Virus Infection

---

Silvina Elena Gutiérrez, Eduardo Néstor Esteban,  
Claudia María Lützelschwab and  
Marcela Alicia Juliarena

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/65804>

---

## Abstract

The major histocompatibility complex (MHC) is a polymorphic gene cluster of about 150 genes, present in all vertebrates. Many of these genes contribute to immunity. Particularly, MHC-encoded class I and class II molecules, which are typically highly polymorphic and polygenic, are central in defining the specificity of the adaptive immune response. Among the diversity of genes associated with disease resistance, MHC genes are particularly interesting as they are associated with resistance and susceptibility to a wide range of diseases, some of which produce important economic losses in livestock. Enzootic bovine leukosis is an infectious disease caused by the retrovirus bovine leukemia virus (BLV), with an important economic impact, mainly in dairy herds. In this chapter, MHC-associated genetic resistance to BLV is revised. Certain alleles of the bovine MHC (*BoLA*) class II locus have been found strongly associated with resistance to viral dissemination. Genetic selection of resistant animals emerges as a natural strategy for the control of infectious diseases, especially when there is no other alternative of control or prevention, as vaccines. Founded on this knowledge, a BLV control program based on selection of genetically resistant cattle was designed. The proof of concept indicates that this strategy is feasible to implement in dairy herds.

**Keywords:** major histocompatibility complex, BoLA, genetic resistance, infectious diseases, bovine leukemia virus, control, veterinary, livestock

## 1. Introduction

The immune system has evolved in vertebrates to protect them from invading pathogens. To attain this objective, it compromises an enormous variety of cells and molecules that interact with each other in a complex network to recognize, counteract, and, if properly regulated, eliminate the pathogen. The major histocompatibility complex (MHC), which is found to occur in all mammalian species, plays a central role in the development and function of the immune system. Genes encoding the MHC are highly polymorphic, and numerous associations between allelic variants and immune responsiveness and disease resistance are well documented. Hence, the MHC genes are attractive as candidate genes involved in susceptibility/resistance to various diseases.

Breeding for improved disease resistance has emerged as a major challenge for animal geneticists. The benefits of successfully improving the resistance of animals to an infectious disease are manifold, including animal welfare, increased efficiency and productivity, and hence a reduced environmental footprint, reduced reliance on other disease-control measures, and improved public perception [1].

Enzootic bovine leucosis is an endemic disease in many countries, causing important economic impact in the dairy industry. The fine characterization of the resistance phenotype, the strong association between certain MHC class II alleles with resistance, and the absence of preventive or therapeutic measures against the disease make the genetic selection of resistant animals a feasible approach to control bovine leukemia virus (BLV) infection.

## 2. The major histocompatibility complex

Vertebrates have the capacity to recognize, destroy, and develop immunological memory to invading microorganisms through the activation of cells and molecules of their immune system. In order to achieve these ends, the two arms of the immune system (i.e., the innate and the acquired immunity) have to interact with each other. The innate immune system comprises mainly cells from the myeloid lineage that recognize common structures on a broad spectra of microorganisms, known as the pathogen-associated molecular patterns (PAMPs) through their pathogen recognition receptors (PRRs). Some of these cells, like macrophages and dendritic cells, are also involved in the activation of the adaptive immune system, by capturing and processing the antigens and acting as antigen presenting cells (APC) for T lymphocytes.

The effector cells of the adaptive immune system, consisting of B and both helper T lymphocytes (LTH) and cytotoxic T lymphocytes (CTL), recognize a very large variety of self and nonself antigens in a more specific manner by their antigen receptors (BCR and TCR, respectively). The BCR can bind directly to free or soluble native antigen, while the TCR requires the protein antigen to be processed into small peptides. These peptides have to be associated within the endoplasmic reticulum with molecules encoded by a single genetic locus containing many polymorphic genes, the MHC. The assembled MHC molecule (mMHC)-peptide

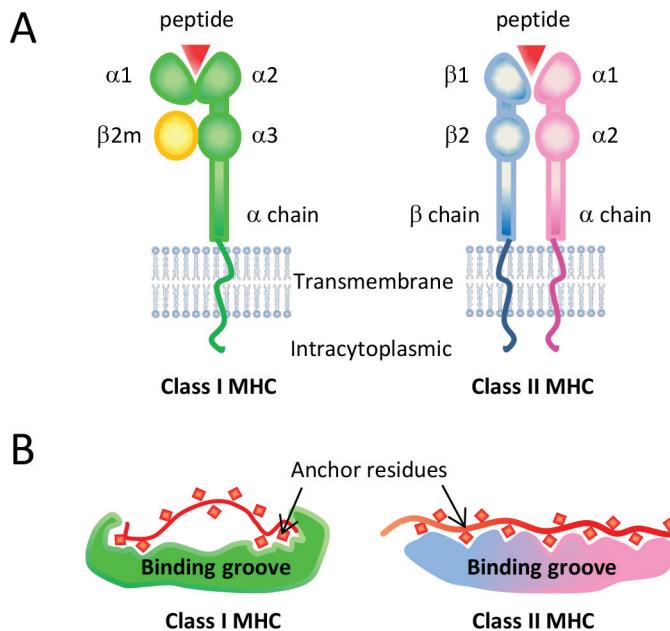
complex is then transported to and expressed on the surface of the APCs, where the antigen-loaded mMHC interacts with the TCR and initiates the activation of T lymphocytes [2]. The CTL and the LTH recognize antigen in the context of two different mMHCs: class I and class II, respectively. The class II mMHCs have a restricted expression, mainly on professional APCs, and expose peptides mainly derived from captured extracellular antigens for the recognition by the LTH. The class I mMHCs are displayed on the surface of every nucleated cell of the organism, presenting peptides essentially originated from intracellular proteins (i.e., intracellular microorganisms or cell-derived proteins) to be recognized by the CTLs. The class I mMHC does not only play a fundamental role in the recognition of foreign intracellular antigens but also in inducing self-tolerance and alloreactive immune responses.

## 2.1. Structure of class I and class II MHC molecules

Most of the current knowledge about these molecules arose from pioneer studies on the rejection of normal and malignant transplanted tissues in mice and rabbits [3–6]. The evidence produced by these studies indicated that the destruction of the grafted tissue was determined upon the existence of inherited antigenic differences between transplant and host, leading to the discovery of the mMHC and, later on, their genetic complexity.

The mMHCs have different domain organizations but similar structure. The variations are concentrated in three to four discrete hypervariable regions in the extracellular domains, while the rest of the molecule is highly conserved. The X-ray crystallography of the proteins demonstrated that the class I mMHC is a heterodimer consisting of a transmembrane  $\alpha$  chain non-covalently linked to a small non-transmembrane chain, called  $\beta 2$ -microglobulin ( $\beta 2m$ ). Besides a transmembrane and an intracellular region, the  $\alpha$  chain has three extracellular globular domains ( $\alpha 1$ ,  $\alpha 2$ , and  $\alpha 3$ ). The  $\alpha 1$  and  $\alpha 2$  domains of the  $\alpha$  chain form a groove that accommodates an 8- to 10-mer antigenic peptide [7, 8]. The class II mMHC is a heterodimer composed of an  $\alpha$  and a  $\beta$  chain, both with an intracellular, a transmembrane, and two extracellular domains (**Figure 1A**). The pairing of the  $\alpha 1$  and  $\beta 1$  domains form an antigen-binding groove with open ends, allowing the allocation for a larger peptide (14-mer or more) extending out of both sides of the groove [9] (**Figure 1B**).

The polygenic and polymorphic features of the MHC grant these molecules with an enormous capability for antigen presentation. A single individual co-expresses several mMHCs from a large pool of alleles within a population. Moreover, each mMHC molecule can associate with a great amount of similar peptides, expanding even more the breadth of MHC-regulated immune responses to pathogens. The polymorphic residues in the mMHC antigen-binding groove are responsible for the different peptide specificities of the different alleles. The class I molecules require an allele-specific peptide length and a defined peptide motif that includes two anchor amino acids or residues with closely related side chains, which interact with both ends of the antigen-binding site of that particular mMHC [10]. The class II mMHCs have less stringency for size but also have allele-specific peptide motifs (or anchor residues) that reach into pockets within and on the sides of the groove of the class II mMHC [11] (**Figure 1B**). Moreover, conserved residues within the class II peptide-binding groove induce a conformational change, forcing bound peptides into a twisted configuration and exposing sites for external interactions [12].



**Figure 1.** Structure of MHC molecules and its binding sites. Panel A: schematic representation of class I and class II mMHCs. Panel B. MHC peptide-binding sites. Amino acid positioning within the peptide-binding grooves of class I and class II MHC proteins is shown. Squares represent individual amino acids of the antigenic peptides binding in the groove of each class I and class II mMHC.

## 2.2. Organization of the Major Histocompatibility Complex

Since its first description in mice and humans [5, 13], the study of the MHC genetic architecture has expanded substantially, with the discovery and characterization of many class I and class II genes in different vertebrate species, except for jawless fish [14]. The collective name given to the proteins encoded by MHC genes depends on the species, except for mice and chickens, in which they were first described as transplantation antigens and maintain their original nomenclature H-2 and B, respectively. Hence, in humans, they are called human leukocyte antigen (HLA); in swine, SLA; in ovine, OLA; in equine, ELA; in dogs, DLA; in bovine, BoLA; and so on. The genetic structure of the MHC is best known for HLA and is relatively conserved among other mammalian species [15].

The HLA complex covers about 4 Mb of the short arm of chromosome 6 and contains three major regions with the confirmed presence of more than 260 loci, including over 160 protein-coding genes [16, 17].

The HLA locus is divided into three closely linked regions: class I, class II, and class III. The first two regions contain genes that control the specific immune response (so-called “classical” MHC genes), and the class III region, enclosing about 75 genes, encodes a variety of different proteins, some related to the innate immunity [18].

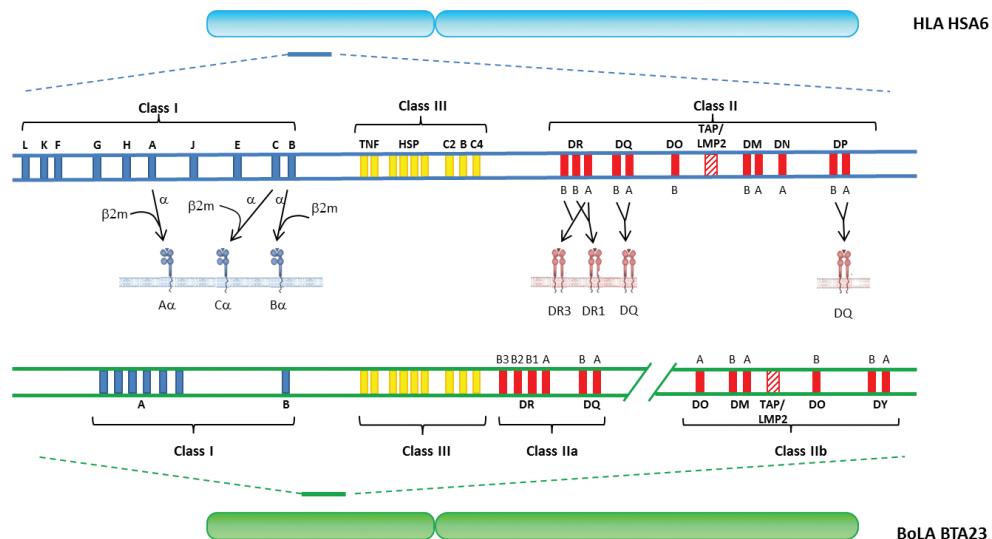
The class I region contains three classical genes, *HLA-A*, *HLA-B*, and *HLA-C*, and three non-classical genes: *HLA-E*, *HLA-F*, and *HLA-G* [19] (Figure 2). Each of the classical HLA is a single functional gene, encoding a class I mMHC  $\alpha$  chain. The gene coding for the  $\beta$ 2m is located outside the MHC, on chromosome 15. *HLA-H*, *HLA-J*, *HLA-K*, and *HLA-L* are nonfunctional pseudogenes, closely related in nucleotide sequence to the class I functional genes [18].

The class II HLA cluster comprises three classical class II genes: *HLA-DP*, *HLA-DQ*, and *HLA-DR*, each encoding one  $\alpha$  and one or two  $\beta$  chains; three nonclassical, non-polymorphic class II genes; *HLA-DM*, *HLA-DN*, and *HLA-DO*; and some pseudogenes [19] (Figure 2). These nonclassical class II genes are not expressed on the cell surface, but form heterotetrameric complexes involved in catalytic peptide exchange and loading onto classical class II molecules [20, 21].

The class III region is located between the class I and class II regions and contains genes coding for molecules with diverse function. Among the most prominent are the complement factor genes coding for factors C2, C4, and B; genes coding for cytokines belonging to the tumor necrosis factor (TNF) superfamily, TNF- $\alpha$ , lymphotoxin- $\alpha$  and lymphotoxin- $\beta$ , which are involved in various inflammatory pathways; heat shock protein genes; and many other genes encoding proteins not related to the immune system [19].

### 2.3. The major histocompatibility complex in cattle

In cattle, the first evidence for the existence of a MHC system was found by lymphocyte immunizations and by the generation of monospecific antilymphocyte antisera against skin grafts, followed by studies on the inheritance of the antigens they detected [22, 23]. As in the human, the *BoLA* locus is highly complex and contains about 154 predicted functional genes spanning about 4 centimorgan on chromosome 23 [24]. **Figure 2** shows the *BoLA* organization



**Figure 2.** An abridge map of the genetic organization of HLA and BoLA.

compared to its human counterpart. Compared to the HLA system, the BoLA system differs in gene arrangement without compromising its functions. Class I genes are clustered in two regions: *BoLA-A* and *BoLA-B*. Only the *A* locus seems to be functional with at least six putative classical class I genes, named *genes 1–6* [25]. To date, 96 *BoLA-A* alleles are listed on the Immuno Polymorphism Database (IPD)-MHC database [26], and 29 different haplotypes have been identified, expressing between one and three combinations of these classical class I genes [27, 28]. The nonclassical BoLA class I genes include *NC1*, *NC2*, and *NC3*, encoding molecules with a relevant role during reproduction in dairy cows [29].

The main difference between BoLA and HLA gene organization is that in cattle the class II gene cluster is divided into two subregions, separated about 15 cM apart from each other: the class IIa, located near the class I/III regions, and the class IIb, resulting from a transposition, located close to the centromere on autosome 23 [30, 31]. This feature is shared by various ruminant species [32].

The subregion IIa incorporates the gene clusters *DR* and *DQ* but lack a *DP* gene [27]. This subregion expresses one DR molecule and one or two DQ molecules per haplotype [33]. Bovines have one monomorphic *DRA* gene. By contrast, there are three genes that encode for the  $\beta$  chain of the DR (*DRB*) molecule of which *DRB1* is a pseudogene and *DRB2* is poorly expressed, leaving the *DRB3* locus as the most polymorphic and strongly expressed gene from this group. To date, 130 alleles have been identified using different approaches in various breeds of cattle (listed in IPD-MHC database) [34].

The *DQ* cluster comprises five *DQA* (*DQA 1–5*) and five *DQB* (*DQB 1–5*) genes, which have arisen from gene duplication. From this cluster, at least 61 *BoLA DQA* and 81 *BoLA DQB* alleles have been described and listed in the IPD-MHC database, and 56 different haplotypes were described in Japanese Black and Holstein cattle [35]. The high number of different alleles, both on the *DR* and *DQ* loci, suggests that DR and DQ molecules complement each other for the presentation of a broad spectrum of antigens in cattle.

The BoLA class IIb locus is divided into two regions. The region known as “extended class II” region contains some genes involved in antigen processing and transport, i.e., LMP complex (low molecular mass polypeptide, *LMP2*, *LMP7*) and *TAP* genes (transporter associated with antigen processing: *TAP2.1*, *TAP1*, and *TAP2*) and also some non-MHC genes like *H2B* (histone H2B-like), among others [36–38]. The second region is known as “classical class II” region and encloses genes of unknown function and unique to ruminants: *DSB* (*DR $\beta$ -like*), *DYA*, and *DYB*. The *DYA* and *DYB* genes encode for proteins of 253 and 259 amino acids, respectively. DY molecule has been shown to be expressed by a subpopulation of afferent lymph dendritic cells, suggesting its involvement in the prominent antigen processing and presentation capability of these cells [39–41].

## 2.4. Association of the major histocompatibility complex to disease susceptibility and resistance

Although age, stress, and physiologic status are important factors influencing the outcome of infection, evidence for genetic control has been observed in many animal species. Based on data registered in the Domestic Animal Diversity Information System (DAD-IS) 2015 report [42], 483 mammalian breed populations are recorded as having resistance or tolerance to

specific diseases or parasites, among which 236 correspond to breeds of cattle, 94 breeds of sheep, 56 breeds of chicken, 54 breeds of goats, and 36 breeds of pigs.

Not unexpectedly, the high degree of genetic polymorphism in the MHC has been associated with health status, vaccine responsiveness, and production traits in cattle [43–47]. Examples of some studies for which there is documented evidence of MHC association with resistance or susceptibility to disease include mastitis in cattle and sheep [44, 48], tick-borne disease [49–51], dermatophilosis in cattle [52], enzootic bovine leucosis in cattle and sheep [53–56], neosporosis in cattle [57], theileriosis in cattle [58], gastrointestinal parasites in sheep [59, 60], diarrhea in pigs [61, 62], Marek's disease in chicken [63, 64], coccidiosis in chicken [65], and coronavirus resistance in chickens [66], among others.

### 3. Enzootic bovine leukosis and the bovine leukemia virus

Enzootic bovine leukosis is one of the most frequent neoplastic diseases of cattle caused by an exogenous retrovirus designated BLV. BLV is the type species of the genus *Deltaretrovirus* in the Retroviridae family. This genus also includes pathogenic viruses from human and nonhuman primates (human and simian T-cell leukemia viruses) that share biological and molecular characteristics with BLV [67].

BLV infection is globally distributed in cattle-raising countries. An assessment of BLV infection in US dairy operations in 2007 showed that 83.9 % of them were seropositive for BLV [68]. A national study of BLV infection in Canada in 1980 showed that 40 % of its dairy herds and 11 % of its beef herds were infected [69]. On the other hand, BLV control programs have been established in member countries of the European Union since the 1980s, resulting in seroprevalence between 0.5 and 1.5 % in some countries, while others such as Belgium, Denmark, Germany, Estonia, Spain, France, Ireland, Austria, Finland, Sweden, the United Kingdom, and few others are considered officially free by the European Community [70–72].

In natural conditions BLV only infects cattle, zebras, buffalos, and capybaras, but other species such as sheep, goats, and rabbits can be experimentally infected [73]. Although both beef and dairy breeds are equally susceptible to BLV infection [74], the impact is higher in dairy herds, mainly because of differential management practices.

The major target of the virus is the B lymphocyte [75]. Although evidence of infection of other peripheral blood cell subpopulations has been reported [76], these results have not been confirmed by others. Soon after infection of a cell, the viral RNA is copied into DNA by the virus-encoded reverse transcriptase. The provirus then integrates into the cellular DNA at random sites, and the infection persists for the whole life of the animal, despite the presence of neutralizing and other antiviral antibodies.

#### 3.1. Pathological and clinical features associated to BLV

BLV infection is characterized by the “iceberg principle,” typical of many viral diseases. While the majority (approximately 70 %) of infected cattle remain asymptomatic, one third of infected cattle develop a permanent increase in the number of B lymphocytes termed persistent

lymphocytosis (PL), which is considered a benign condition. The tip of the iceberg is represented by those animals that develop the neoplastic disease, which is usually less than 5 % of the infected cattle. The accumulation of transformed lymphocytes in one or more organs after a long latency period of 1–8 years leads to a multicentric lymphosarcoma. This condition is typically observed in cattle older than 3 years of age. In two thirds of the animals, the development of tumors is preceded by a phase of PL. Lesions can be localized in almost any organ, but the abomasum, heart, visceral and peripheral lymph nodes, spleen, uterus, and kidneys are most frequently affected. Lesions can be observed as white firm tumor masses or as a diffuse tissue infiltrate in any organ. Clinical signs are variable and depend on the affected organ, the speed of growth of tumors, and the degree of dissemination of the neoplastic process. In most cases the course of the illness is subacute to chronic, initiated by a marked loss of weight and appetite, and weakness. Clinical signs most often observed are decreased milk production, lymphadenopathy, and posterior paresis. Once the clinical signs of the illness are evident, the course is rapid and invariably culminate in death [77].

### 3.2. Transmission of BLV and economic impact of the infection

As cell-free virus is rarely detected *in vivo*, most susceptible cattle become infected by exposure to infected lymphocytes. Vertical transmission may occur *in utero* but is infrequent. The main biologic fluids that contain sufficient infected lymphocytes to transmit the infection are the blood, colostrum, and milk. Other fluids such as saliva, semen, urine, and nasal secretions, while potentially infectious, have not been demonstrated to transmit the infection in natural conditions [78].

Under general or standard production conditions, the risk of horizontal transmission is augmented by management practices or procedures involving blood transfer such as gouge dehorning, ear tagging or tattooing, blood extraction, and rectal palpation, using shared or not properly disinfected instruments [78]. This risk is augmented when contaminating blood comes from cattle with persistent lymphocytosis [79]. The use of natural service (i.e., bulls) to breed heifers was also identified as a risk factor for augmented prevalence of BLV infection compared to artificial insemination [80].

Evidence has been reported on the role of bloodsucking insects, such as stable flies, horn flies, and tabanids in the transmission of BLV [81]. Furthermore, the lack of insect control program has been recognized as a risk factor for BLV infection [82, 83]. In warm regions, the animals may be exposed to a high density of hematophagous insects that continuously feed on them; hence, the control of bloodsucking insects by pesticides has been reported to prevent the transmission in a model farm [84]. Both colostrum and milk from BLV-positive cows contain infected lymphocytes, and evidence exists for the transmission of BLV to calves by feeding bulk milk, a common practice in dairy herds [78, 85]. The rate of transmission attributable to this route has been estimated to be 6–16 % under natural conditions. On the other hand, feeding colostrum from infected dams, which contains high titers of antiviral antibodies, seems to have a protective role, being the susceptibility of calves dependent on the presence of specific antibodies obtained from the dam's colostrums and the age of the calf [78, 86, 87].

Enzootic bovine leucosis causes significant economic losses. The most obvious economic losses are due to culling or death due to lymphosarcoma, shortening of lifespan, and loss of

production potential. Other indirect losses are related to the costs of control and eradication programs and restrictions in the international trade of cattle and their by-products. Annual economic losses to the US dairy industry associated with BLV are estimated to be \$285 million for producers and \$240 million for consumers [80]. The effects of subclinical BLV infection on milk production, reproductive performance, longevity, and culling rate are variable. Using data from the National Animal Health Monitoring System's 1996 dairy herd study, it was found that herds with test-positive cows produced 218 kg less milk per cow/per year than those with no test-positive cows [88].

There is no treatment for BLV infection or its associated disease. The possibility of a vaccine for protection against BLV has been explored (reviewed in [89]). A BLV vaccine would have to be noninfectious and non-oncogenic and should not interfere with the serological tests commonly used to detect infection [77].

### 3.3. Phenotypes associated with resistance and susceptibility in BLV infection

Studies from our laboratory have gone further into the characterization of BLV-infected, hematologically normal cattle (i.e., those animals that do not develop PL). Based on the proviral load in the peripheral blood and antibodies against BLV major proteins, we could describe two defined phenotypes in BLV-infected cattle. Proviral load was determined in DNA from peripheral blood leukocytes by a very sensitive nested PCR or a semiquantitative PCR [90], while antibody titers were measured against the BLV major antigens by ELISA [90, 91].

One group of animals is characterized by high proviral load (HPL) in peripheral blood (>100,000 BLV proviral copies/µg of DNA) and high antibody titers against the envelope protein of BLV of 51 kDa (BLVgp51). On the other hand, the remaining non-PL animals harbor an exiguous number of infected lymphocytes in the peripheral blood, almost undetectable by the molecular methods currently used (PCR and real-time PCR); we have termed this group low proviral load (LPL) cattle. These cattle develop low titers of antiviral antibodies against BLVgp51, while antibodies against BLVp24, the main core protein of BLV, are undetectable in the majority of LPL cattle or developed at very low titers. The follow-up of HPL and LPL cattle showed that each phenotype was maintained throughout, at least, a 1.5-year period. The characterization of PL cattle in terms of proviral load and titers of antibodies against BLV showed no significant differences with non-PL HPL cattle. Cattle with LPL profile represented about 60 % of hematologically normal cattle and 40 % of all BLV-infected cattle [92]. **Table 1** shows the main parameters that characterize both HPL and LPL infection phenotypes.

As virtually all cattle infected with BLV will continuously have antibodies against the virion proteins in their serum, serologic tests are commonly used for diagnosis of BLV infection in cattle > 6 months old. PL cattle and non-PL HPL cattle have high titers of antibodies against the major viral antigens, and then, they are easily classified as positive by means of the majority of serological tests in the market (mainly ELISAs or immunodiffusion tests). On the other hand, antibody titers against BLVgp51 are lower in most LPL animals compared to HPL cattle, and antibodies against BLVp24 are at very low titers or undetectable in most serum samples

BLV infection phenotype	Hematological status	Proviral load <sup>a</sup>	Antibody titer	
			BLVgp51	BLVp24
HPL	PL	≥100,000	400 to ≥ 6400	50–800
HPL	Non PL	≥100,000	400 to ≥ 6400	Seronegative to 400
LPL	Non PL	≤100	2–1600	Seronegative to 50

Proviral load was determined by semiquantitative PCR in DNA extracted from the peripheral blood. Antibody titers against BLVgp51 and BLV p24 were determined in plasma samples by ELISA.

<sup>a</sup> Proviral copies/μg DNA

Adapted from Juliarena et al. [92]

**Table 1.** Characterization of HPL and LPL phenotypes in terms of proviral load in the peripheral blood, hematological status, and antibody titer against BLVgp51 and BLVp24.

from LPL cattle. Hence, sensitivity of serological tests is decisive in their capacity to detect LPL animals. We have consistently detected LPL animals with a highly sensitive blocking ELISA developed in our lab, designated ELISA 108. This method detects antibodies against an immunodominant conformational epitope of the BLVgp51 at low titers. Other serological methods using BLVgp51 as antigen may also detect BLV-infected LPL animals, provided that they are sensitive enough. This is not the case of agar gel immunodiffusion (AGID) test commercialized in Argentina. On the other hand, serological tests using BLVp24 as antigen would not be able to detect LPL cattle in most cases.

We inoculated sheep with blood from LPL cattle to give further evidence of infection. This was an easily available alternative for the amplification and study of the BLV strain [93]. As the number of infected lymphocytes in peripheral blood from LPL cattle is extremely low, it was necessary to inoculate a large volume of blood to obtain the minimal quantity of infected lymphocytes to transmit the infection. The minimal dose necessary to infect sheep is about 926 BLV-infected lymphocytes [94]. One microliter of blood from a HPL cow was enough to infect lambs. In contrast, blood from LPL cattle also infected lambs, but 100 ml of blood was necessary to infect the ovine host. In order to determine if differences in the observed phenotypes could be attributable to differences at the nucleotide level between strains, a fragment of the *env* gene was amplified from infected lambs by nested PCR and sequenced. No mutations in the *env* gene that could be attributable to the passage from cow to lamb were detected. When comparing a 400-bp sequence from the *env* gene from six LPL and six HPL strains, no mutations were found that could be associated with any particular phenotype, strengthening the hypothesis that the development of each phenotype could be associated primarily with some genetic or epigenetic property of the host [95].

As LPL cattle maintained their phenotype for prolonged periods of time, without developing any hematologic or pathologic condition, it is proposed that these animals are naturally resistant to BLV replication.

### 3.4. Association of BLV phenotypes to polymorphisms at the major histocompatibility complex

Early observations on the aggregation of lymphoma/leukemia or PL in certain BLV-infected families, but not in others, suggested that host genetic factors were involved in the development

of these conditions [96]. There were also indications that host genetic factors influenced the susceptibility to one condition or the other independently [97].

Studies on the relationship between BLV infection and bovine MHC (BoLA) revealed an association between serologically determined *BoLA-A* class I antigens and resistance and susceptibility to B-cell lymphocytosis in Shorthorn cattle [98], Holstein [99, 100], and other breeds [101]. However, these associations were relatively weak at the population level, and different *BoLA-A* alleles had significant effects in different breeds [102]. Subsequently, it was shown that resistance and susceptibility to PL map more closely to MHC class II *BoLA-DRB3* gene than to *BoLA-A* locus [56, 103]. A peptide motif named ER, which is present in *BoLA-DRB3* alleles \*11, \*23, and \*28, was associated with resistance to PL in BLV-infected cattle. Resistance appears to be dependent upon the presence of the polar amino acids Glu-Arg at positions 70–71 within a highly polymorphic segment of the peptide-binding region [56]. This contributes to the putative peptide-binding specificity of the molecule at the  $\beta 1$  domain [11]. It is believed that the allelic differences influence the binding and orientation of viral peptides [104], thus determining allele specificity differences in the spectrum of peptides presented to the immune system. These differences may have important consequences for infection resistance [105]. Furthermore, it was shown that *BoLA-DRB3* alleles encoding Glu, Arg, and Val at positions 74, 77, and 78 of the *BoLA-DRB3* chain, respectively, might be related to tumor development resistance [106]. Moreover, evidence was presented that BLV-infected cattle selected for the presence of the *DRB3\*11* allele carry less infected lymphocytes than other infected animals [107].

Based on these previous reports found in literature about the association of *BoLA* polymorphisms with PL development or the number of BLV-infected lymphocytes in peripheral blood up to 1998, we decided to evaluate the association between the BLV infection phenotypes described above and *BoLA-DRB3* genotype. Various molecular methods were used for the genotyping of *BoLA-DRB3* exon 2 (the only one expressed from *BoLA-DRB3* gene). In a first instance, we used the PCR-restriction fragment length polymorphism (PCR-RFLP) method, which can differentiate only 57 different alleles (designated by the nomenclature based on this method). The PCR-sequence specific oligonucleotides polymorphisms (PCR-SSOP) is an alternative method which is more specific, allowing the identification of 104 alleles from the 130 described by sequencing, and defines the alleles according to the nomenclature adopted by the International Society for Animal Genetics [34].

Xu et al. [56] described a correlation between *BoLA-DRB3\*11* allele and PL in Holstein cattle. The allele \*11 (identifiable by PCR-RFLP) can be differentiated into two variants: *BoLA-DRB3\*0901* and *BoLA-DRB3\*0902* by using PCR-SSOP or by sequencing. By studying the genotype of 230 BLV-infected cattle belonging to seven dairy herds, we found that allele *BoLA-DRB3\*11* was significantly associated with the LPL phenotype (odds ratio (OR) = 5.82;  $p < 0.0001$ ); however, the subtype *BoLA-DRB3\*0902*, which is the most prevalent subtype of allele \*11 in our population, showed a stronger association with the LPL phenotype (OR = 8.24;  $p < 0.0001$ ) than allele \*11 itself. Allele *BoLA-DRB3\*1701* also showed significant association with LPL profile (OR = 3.46;  $p < 0.0055$ ). The HPL phenotype was significantly associated (OR = 0.36;  $p < 0.0005$ ) with only one allele: *BoLA-DRB3\*1501* (allele \*16 as determined by PCR-RFLP). According to these associations, the DRB3 alleles were assigned to three categories: resistant (R) if the allele

was associated with LPL phenotype, susceptible (S) if the allele was associated with HPL phenotype, and neutral (N) if the allele was not associated with any phenotype [53].

It is concluded that the host genetic background influences BLV infection phenotype development and that allele *BoLA-DRB3\*0902* appears to be, up to now, the best candidate marker of BLV resistance in Argentinean Holstein cattle. The penetrance of *BoLA-DRB\*0902* allele for the LPL phenotype is notably high compared to the penetrance of MHC alleles in other systems which, in general, is much lower (about 10%). More than 80 % of cattle carrying the *BoLA-DRB3\*0902* allele develop LPL when naturally or experimentally infected with BLV [53, 108, 109]. However, only one third of LPL cattle harbors the *BoLA-DRB3\*0902* allele [53]. This finding suggests that other genetic or epigenetic factors might be involved in the regulation of BLV proviral load.

### 3.5. Influence of resistance-associated alleles on specific immune response and infectivity of LPL cattle

Cattle with LPL phenotype were grouped according to the presence of alleles associated to the LPL phenotype (i.e., R alleles \*0902 and \*1701) in their genotype. **Table 2** shows the average anti-BLVgp51 antibody titer for each group and the resulting classification of cattle according to the level of antibody response against the main envelope glycoprotein of BLV. The R alleles *BoLA-DRB3\*0902* and *BoLA-DRB3\*1701* were significantly associated with low antibody titers against both BLVgp51 and BLVp24 [110]. Important differences were observed in the infectivity of the blood from cattle in each group, as determined by the sheep bioassay. While all lambs inoculated with blood from LPL cattle not harboring any allele R or harboring one R allele (\*0902 or \*1701) in heterozygosis were infected, only one from six lambs inoculated with 100 ml of blood from LPL cattle with R/R genotype acquired the infection. From these results, it seems that LPL cattle carrying two R alleles have an increased ability to restrict BLV replication [108].

It is unlikely that BLV-infected LPL cattle would be a source of infection for BLV-free animals, as large volumes of blood are never exchanged between animals under usual management practices and natural breeding in dairy farms.

<i>BoLA-DRB3</i> genotype	n	BLVgp51 antibody titer <sup>a</sup>	Classification <sup>b</sup>	Infectivity <sup>c</sup>
Without allele R	54	662.3	Medium or high responders	6/6
Allele R in heterozygosis	34	344.3	Medium or low responders	4/4
Allele R in homozygosis	9	275.0	Extremely low responders	1/6

<sup>a</sup> Average

<sup>b</sup> Classification was done according to the level of antibody titer against BLVgp51.

<sup>c</sup> Infectivity was assessed by inoculating 100 ml of blood in lambs. Number of lambs infected/number of lambs inoculated

**Table 2.** Grouping of BLV-infected cattle with LPL phenotype, according to the presence of alleles associated with resistance (R) in their genotype.

### 3.6. Strategies to control BLV infection

BLV control and eradication programs based on culling infected animals have been accepted and implemented successfully in several western European countries. In dairy farms with low prevalence of BLV infection, the cost of implementing these programs is less than the total cost caused by BLV infection [111]. However, in herds where the prevalence of infection is high, the feasibility of implementation of these programs depends on the official compensation to producers for the removal of BLV-positive animals. In Argentina, not only there is no official BLV control program, but official apathy since the 1970s has been the best ally for the lush spread of BLV in all dairy regions, including Patagonia. By 1995, the failure of the voluntary control program of BLV proposed by the National Authority in Animal Heath (Resolution 337/94 from SENASA) was the indicator that Argentina had lost the opportunity to reduce the spread of BLV by means of serological testing and culling and to protect the scarce herds and regions that were still free of BLV.

To limit the spread of BLV, infected cattle are often eliminated on the basis of their risk of transmitting the infection to other cattle. Among BLV-infected cattle, those with PL are considered the most efficient transmitters because they harbor a high percentage of infected lymphocytes in peripheral blood [112] and consequently have HPL [107]. Furthermore, although not all PL animals develop lymphosarcoma, in approximately two thirds of cases, neoplasia is preceded by PL. The application of control plans based on hematological tests and elimination of PL animals has failed to control not only BLV infection but also the associated lymphosarcoma. In 1959, Denmark implemented an eradication program based on the occurrence of clinical lymphosarcoma and the identification of PL cattle by using the Bendixen hematological keys [113]. Affected herds were quarantined, and indemnity was offered to induce owners to have slaughtered their entire herd. This herd-slaughter policy was continued until 1982. However, when the serological tests were introduced, some herds which were classified as leucosis-free based on the hematological keys were found to be infected [77]. One of the major causes of this failure was probably due to the presence of HPL animals that do not present PL. These non-PL HPL animals are as efficient as PL cattle to transmit the BLV, and it is likely that the animals that develop lymphosarcoma without a previous stage of PL also belong to this group of cattle. Therefore, the corrective management based in classification of animals according to HPL and LPL phenotypes should be safer than the plan based on traditional hematological classification (PL and non-PL).

## 4. Genetic selection of resistant animals

The genetic selection of resistant animals would be an efficient strategy for controlling the spread of BLV in the host and, therefore, to reduce the occurrence of clinical lymphosarcoma. We have shown that, in the Holstein breed, the *BoLA-DRB3\*0902* allele is the best molecular marker for the selection of resistant animals. The *BoLA-DRB3* gene is highly polymorphic, with 130 alleles described so far. The allele frequency of the \*0902 allele is relatively high (5–10 %) at the population level; therefore, it is possible to easily increase its frequency by directed crossbreeding or artificial insemination. At the herd level, it is essential that these

resistant animals disrupt the transmission of BLV to uninfected animals. Hence, the effectiveness of this procedure would depend on the accuracy of the premise that BLV-infected LPL cattle harboring *BoLA-DRB3\*0902* marker do not transmit the virus to BLV-negative cattle under normal breeding conditions in commercial dairy herds.

#### 4.1. The proof of concept

In order to test this premise, an experiment was carried out in a commercial dairy herd located in a subtropical region of Argentina, having extreme environmental conditions (experimental dairy herd). In the experimental dairy herd, BLV-free and BLV-infected LPL-*BoLA-DRB3\*0902* animals cohabited for 20 months. This period included two summers during which BLV-infected and BLV-uninfected cattle were exposed to heat stress and to a high density and varied population of bloodsucking insects that continuously fed on the animals. As is usual in any commercial dairy herd, the cattle population was dynamic, and necessary replacements were made according to production needs and milk production requirements, which were within expected parameters in the region. Nevertheless, only previously characterized animals (BLV-free or BLV-infected LPL-*BoLA-DRB3\*0902* animals) were introduced to the experimental herd. Normal management practices and production requirements were maintained in the herd. We examined the spread of BLV in the experimental herd and also in other four local commercial herds in which no selection of cattle was made. The incidence rates observed in these four herds were between 0.06 and 0.17 cases per 100 cattle-days. The observed incidence rate explains the increase of prevalence registered in the region from 2002–2003 to 2010–2011 [114, 115]. In contrast, in the experimental herd, no new BLV-infected animals were detected, and as a corollary, the incidence rate was zero. These results indicate that LPL-*BoLA-DRB3\*0902* cattle do indeed disrupt the BLV transmission chain and that selection of cattle carrying this allele represents a promising approach to control the virus [116].

The results presented are the base of a BLV control strategy based on genetic selection of resistant animals, using *BoLA-DRB3\*0902* allele as molecular selection marker. The idea is to gradually replace HPL cattle in high-prevalence herds with *BoLA-DRB3.2\*0902* harboring cattle. The expected increase in herd resistance to BLV would block the transmission of BLV, thereby counteracting the main epidemiological advantage of BLV, i.e., its transmission.

#### 4.2. Influence of resistance and susceptibility-associated alleles on other health and productive traits

Due to the relevant role of MHC genes in the immune response, a potential risk in expanding or segregating *BoLA*-selected populations of cattle is that it might increase susceptibility to other common viruses. A special concern is raised by the strong association found between *BoLA-DRB3\*0902* and *BoLA-DRB3\*1701* and low antibody titer against major BLV structural proteins. This phenomenon, which may depend on host genetic factors, might influence the host response to other viruses requiring, unlike BLV, strong and long-lasting humoral immune response in order to prevent infection. Therefore, we determined the association between antibody titer against three widespread bovine viruses and *BoLA-DRB3* gene polymorphism. These three viruses, the bovine viral diarrhea virus (BVDV), the

bovine herpesvirus type 1 (BHV-1), and the foot and mouth disease virus (FMDV), which are spread in most or several countries all over the world, cause diseases and syndromes that negatively affect the economic performance of the dairy and beef cattle industries. Protection against these three viruses is achieved when high titers of neutralizing antibodies are raised. No association was found between neutralizing antibody titers against FMDV, BVDV, or BHV-1 and polymorphism of the *BoLA-DRB3* gene [110]. Therefore, increasing *BoLA*-selected BLV-resistant cattle or segregating *BoLA* alleles associated with BLV susceptibility would not affect the resistance or susceptibility to BVDV, BHV-1, or FMDV infection. Further studies involving other parameters, such as the persistence or viral load status concerning FMDV, BHV-1, and BVDV, and their association to *BoLA-DRB3* polymorphism, are needed. These parameters would probably represent a better indicator of resistance against the aforementioned viruses.

Preliminary data obtained from endemic areas showed that there is no association between the allele *BoLA-DRB3\*0902* and infection with BVDV, bovine herpesvirus type 4 (BoHV-4), and *Neospora caninum* [117]. A recent study also showed no association of the aforementioned allele with infection by *Mycobacterium bovis* (*M. bovis*), although this observation should be confirmed in a more extensive study. Moreover, it was shown that the ability of cattle carrying resistance-associated marker to control BLV and to progress to LPL phenotype was not altered by *M. bovis* coinfection [118].

Alleles *BoLA-DRB3\*0902* (or \*11 as determined by PCR-RFLP) and *BoLA-DRB3\*1701* (or \*12 as determined by PCR-RFLP) are associated with resistance to intramammary infection and higher production traits [119, 120]. Thus, expanding the population of cattle harboring these alleles in order to control BLV infection would also increase resistance to mastitis. When targeting BLV control by means of genetic selection, animals carrying alleles associated to susceptibility, such as *BoLA-DRB3\*1501*, are not desired. This allele did not show association with antibody titer against FMDV, BHV-1, or BVDV [110] nor it affected productive traits [45]. *BoLA-DRB3\*1501* has been previously associated with high somatic cell score (SCS) in milk [44, 121]. Thus, if we are to decrease the frequency of the *BoLA-DRB3\*1501* allele in order to control BLV infection, the SCS of the cattle population would additionally be reduced. It has been reported that BLV infection is frequently associated with an increase in SCS [122–124]. It is conceivable that this association derives from host genetic susceptibility to both conditions, probably conferred by *BoLA-DRB3\*1501* allele.

Up to date, the presence of *BoLA-DRB3\*0902* allele in the genotype of Holstein cattle neither has it been associated with susceptibility to any infectious disease nor with negative effects on production or reproductive traits.

#### 4.3. Concluding remarks

We have proposed a BLV control and eradication program based on genetic selection. This plan is based on two main premises: (1) the development of LPL phenotype (BLV-resistant cattle) can be predicted by a marker (*BoLA-DRB3\*0902*) that is transmitted to the progeny by controlled crossbreeding, and (2) BLV-infected cattle with LPL phenotype do not transmit the virus.

The *BoLA-DRB3\*0902* heterozygous animals comply these two premises. This marker has a penetrance of 82 % and its pattern of inheritance is codominant. *BoLA-DRB3\*0902* animals with LPL phenotype did not transmit the infection at herd level in natural conditions, and under experimental conditions, it was required to inoculate a large volume of blood to infect sheep. Finally, up to the date, the *BoLA-DRB3\*0902* allele has not been associated with susceptibility to other infectious agents, nor has it affected production or reproduction traits.

The BLV control and eradication program has been designed in two steps: as the first step, HPL cattle should be replaced by cattle harboring the *BoLA-DRB3\*0902* marker. Once culling of HPL animals has been completed, as the second step, LPL cattle should be replaced by BLV-negative cattle. The selected animals for the replacement in the first step should carry the *BoLA-DRB3\*0902* allele in heterozygosis with other alleles not associated with any BLV phenotype [53]. Finally, it should be mentioned that a potential risk exists in expanding or segregating *BoLA* genotype-selected populations, which might increase susceptibility to other infectious agents.

## 5. Conclusions

Among the diversity of genes associated to disease resistance, MHC genes are particularly interesting as they are associated with resistance and susceptibility to a wide range of diseases. Some of these diseases produce important economic losses in livestock. Genetic selection of resistant animals emerges as a natural strategy for the control of infectious diseases, especially when there is no other alternative of control or prevention, as vaccines.

The control of BLV infection based on genetic selection of resistant cattle is presented as a model that may potentially be replicated in other infectious diseases affecting livestock.

## Author details

Silvina Elena Gutiérrez<sup>1\*</sup>, Eduardo Néstor Esteban<sup>1, 3</sup>, Claudia María Lützelschwab<sup>1</sup> and Marcela Alicia Juliarena<sup>1, 2</sup>

\*Address all correspondence to: segutier@vet.unicen.edu.ar

1 Department of Animal Health and Preventive Medicine, Veterinary Research Center of Tandil (CIVETAN), CIC-CONICET, Faculty of Veterinary Science, National University of the Center of Buenos Aires Province, Tandil, Argentina

2 Department of Animal Production, Veterinary Research Center of Tandil (CIVETAN), CIC-CONICET, Faculty of Veterinary Science, National University of the Center of Buenos Aires Province, Tandil, Argentina

3 Director of BIOALPINA Program (GENIAL/COTANA), Colonia Alpina, Santiago del Estero, Argentina

## References

- [1] Bishop SC, Axford RFE, Nicholas FW, Owen JB Principles and Methods. In: Bishop SC, Axford RFE, Nicholas FW, Owen JB, editor. *Breeding for Disease Resistance in Farm Animals*. Wallingford: CAB International; 2010. p. 1-54.
- [2] DeFreitas EC, Chesnut RW, Grey HM, Chiller JM. Macrophage-dependent activation of antigen-specific T cells requires antigen and a soluble monokine. *Journal of Immunology*. 1983;131(1):23-9.
- [3] Tyzzer EE. A study of inheritance in mice with reference to their susceptibility to transplantable tumors. *The Journal of Medical Research*. 1909;21(3):519-73.
- [4] Little CC, Tyzzer EE. Further experimental studies on the inheritance of susceptibility to a transplantable tumor, carcinoma (J. W. A.) of the Japanese waltzing mouse. *The Journal of Medical Research*. 1916;33(3):393-453.
- [5] Gorer PA. The genetic and antigenic basis of tumour transplantation. *The Journal of Pathology and Bacteriology*. 1937;44(3):691-7.
- [6] Medawar PB. The behaviour and fate of skin autografts and skin homografts in rabbits: A report to the War Wounds Committee of the Medical Research Council. *Journal of Anatomy*. 1944;78(Pt 5):176-99.
- [7] 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(6139):506-12.
- [8] 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(6139):512-8.
- [9] Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*. 1993;364(6432):33-9.
- [10] Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*. 1991;351(6324):290-6.
- [11] Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature*. 1994;368(6468):215-21.
- [12] Allen PM, Matsueda GR, Evans RJ, Dunbar JB, Jr., Marshall GR, Unanue ER. Identification of the T-cell and la contact residues of a T-cell antigenic epitope. *Nature*. 1987;327(6124):713-5.
- [13] Dausset J. Iso-leuko-antibodies. *Acta Haematologica*. 1958;20(1-4):156-66.
- [14] Flajnik MF, Kasahara M. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity*. 2001;15(3):351-62.

- [15] Vaiman M, Chardon P, Cohen D. DNA polymorphism in the major histocompatibility complex of man and various farm animals. *Animal Genetics*. 1986;17(2):113-33.
- [16] Horton R, Gibson R, Coggill P, Miretti M, Allcock RJ, Almeida J, et al. Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. *Immunogenetics*. 2008;60(1):1-18.
- [17] Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. *Annual Review of Genomics and Human Genetics*. 2013;14:301-23.
- [18] Marsh SGE, Parham P, Barber LD. 3 - The Organization of HLA Genes Within the HLA Complex. *The HLA FactsBook*. London: Academic Press; 2000. p. 7-13.
- [19] Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of Human Genetics*. 2009;54(1):15-39.
- [20] Liljedahl M, Kuwana T, Fung-Leung WP, Jackson MR, Peterson PA, Karlsson L. HLA-DO is a lysosomal resident which requires association with HLA-DM for efficient intracellular transport. *The EMBO Journal*. 1996;15(18):4817-24.
- [21] Alfonso C, Karlsson L. Nonclassical MHC class II molecules. *Annual Review of Immunology*. 2000;18(1):113-42.
- [22] Amorena B, Stone WH. Serologically defined (SD) locus in cattle. *Science*. 1978;201(4351):159-60.
- [23] Spooner RL, Levezel H, Grosclaude F, Oliver RA, Vaiman M. Evidence for a possible major histocompatibility complex (BLA) in cattle. *Journal of Immunogenetics*. 1978;5(5):325-46.
- [24] BGSA C, Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009;324(5926):522-8.
- [25] Birch J, Murphy L, MacHugh ND, Ellis SA. Generation and maintenance of diversity in the cattle MHC class I region. *Immunogenetics*. 2006;58(8):670-9.
- [26] IPD. (Immuno Polymorphism Database). MHC Database. BoLA Nomenclature. Cattle: European Bioinformatics Institute; Available from: [https://www.ebi.ac.uk/cgi-bin/ipd/mhc/view\\_nomenclature.cgi?bola.classi](https://www.ebi.ac.uk/cgi-bin/ipd/mhc/view_nomenclature.cgi?bola.classi).
- [27] Ellis SA, Ballingall KT. Cattle MHC: evolution in action? *Immunological Reviews*. 1999;167:159-68.
- [28] Aida Y, Takeshima S, Baldwin CL, Kaushik AK. Bovine immunogenetics. In: Garrick DJ, Ruvinsky A, editors. *The genetics of Cattle*. 2nd ed. Croydon, UK: CAB International; 2015. p. 153-91.
- [29] Shu L, Peng X, Zhang S, Deng G, Wu Y, He M, et al. Non-classical major histocompatibility complex class makes a crucial contribution to reproduction in the dairy cow. *The Journal of Reproduction and Development*. 2012;58(5):569-75.

- [30] Andersson L, Lunden A, Sigurdardottir S, Davies CJ, Rask L. Linkage relationships in the bovine MHC region. High recombination frequency between class II subregions. *Immunogenetics*. 1988;27(4):273-80.
- [31] McShane RD, Gallagher DS, Newkirk H, Taylor JF, Burzlaff JD, Davis SK, et al. Physical localization and order of genes in the class I region of the bovine MHC. *Animal Genetics*. 2001;32(5):235-9.
- [32] Skow LC, Simples SN, Davis SK, Taylor JF, Huang B, Gallagher DH. Localization of bovine lymphocyte antigen (BoLA) DYA and class I loci to different regions of Chromosome 23. *Mammalian Genome*. 1996;7(5):388-9.
- [33] Takeshima S-N, Aida Y. Structure, function and disease susceptibility of the bovine major histocompatibility complex. *Animal Science Journal*. 2006;77(2):138-50.
- [34] IPD. Immuno Polymorphism Database—MHC Database. BoLA Nomenclature, Cattle, BoLA DRB3: European Bioinformatics Institute; Available from: [https://www.ebi.ac.uk/cgi-bin/ipd/mhc/view\\_nomenclature.cgi?bola.drb3](https://www.ebi.ac.uk/cgi-bin/ipd/mhc/view_nomenclature.cgi?bola.drb3).
- [35] Miyasaka T, Takeshima SN, Sentsui H, Aida Y. Identification and diversity of bovine major histocompatibility complex class II haplotypes in Japanese Black and Holstein cattle in Japan. *Journal of Dairy Science*. 2012;95(1):420-31.
- [36] Shalhev D, Da Y, Beever JE, van Eijk MJ, Ma R, Lewin HA, et al. Genetic mapping of the LMP2 proteasome subunit gene to the BoLA class IIb region. *Immunogenetics*. 1995;41(1):44-6.
- [37] Behl JD, Verma NK, Tyagi N, Mishra P, Behl R, Joshi BK. The major histocompatibility complex in bovines: a review. *ISRN Veterinary Science*. 2012;2012:872710.
- [38] Childers CP, Newkirk HL, Honeycutt DA, Ramlachan N, Muzney DM, Sodergren E, et al. Comparative analysis of the bovine MHC class IIb sequence identifies inversion breakpoints and three unexpected genes. *Animal Genetics*. 2006;37(2):121-9.
- [39] Ballingall KT, Ellis SA, MacHugh ND, Archibald SD, McKeever DJ. The DY genes of the cattle MHC: expression and comparative analysis of an unusual class II MHC gene pair. *Immunogenetics*. 2004;55(11):748-55.
- [40] Ballingall KT, McKeever DJ. Conservation of promoter, coding and intronic regions of the non-classical MHC class II DYA gene suggests evolution under functional constraints. *Animal Genetics*. 2005;36(3):237-9.
- [41] Ballingall K, MacHugh N, Taracha E, Mertens B, McKeever D. Transcription of the unique ruminant class II major histocompatibility complex-DYA and DIB genes in dendritic cells. *European Journal of Immunology*. 2001;31(1):82-6.
- [42] FAO. Animal Genetic Resources and Resistance to Disease. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO); 2015.
- [43] Weigel KA, Freeman AE, Kehrli ME, Jr., Stear MJ, Kelley DH. Association of class I bovine lymphocyte antigen complex alleles with health and production traits in dairy cattle. *Journal of Dairy Science*. 1990;73(9):2538-46.

- [44] Dietz AB, Cohen ND, Timms L, Kehrli ME, Jr. Bovine lymphocyte antigen class II alleles as risk factors for high somatic cell counts in milk of lactating dairy cows. *Journal of Dairy Science*. 1997;80(2):406-12.
- [45] Sharif S, Mallard BA, Wilkie BN, Sargeant JM, Scott HM, Dekkers JC, et al. Associations of the bovine major histocompatibility complex DRB3 (BoLA-DRB3) with production traits in Canadian dairy cattle. *Animal Genetics*. 1999;30(2):157-60.
- [46] Glass EJ. Genetic variation and responses to vaccines. *Animal Health Research Reviews/Conference of Research Workers in Animal Diseases*. 2004;5(2):197-208.
- [47] Glass EJ, Baxter R, Leach RJ, Jann OC. Genes controlling vaccine responses and disease resistance to respiratory viral pathogens in cattle. *Veterinary Immunology and Immunopathology*. 2012;148(1-2):90-9.
- [48] Raadsma HW, Gray GD, Woolaston RR. Breeding for disease resistance in Merino sheep in Australia. *Revue Scientifique et Technique*. 1998;17(1):315-28.
- [49] Duangjinda M, Jindatajak Y, Tipvong W, Sriwarothai J, Pattarajinda V, Katawatin S, et al. Association of BoLA-DRB3 alleles with tick-borne disease tolerance in dairy cattle in a tropical environment. *Veterinary Parasitology*. 2013;196(3-4):314-20.
- [50] Shyma KP, Gupta JP, Singh V. Breeding strategies for tick resistance in tropical cattle: a sustainable approach for tick control. *Journal of Parasitic Diseases*. 2015;39(1):1-6.
- [51] Untalan PM, Pruitt JH, Steelman CD. Association of the bovine leukocyte antigen major histocompatibility complex class II DRB3\*4401 allele with host resistance to the Lone Star tick, *Amblyomma americanum*. *Veterinary Parasitology* 2007;145(1-2):190-5.
- [52] Maillard JC, Chantal I, Berthier D, Thevenon S, Sidibe I, Razafindraibe H. Molecular immunogenetics in susceptibility to bovine dermatophilosis: a candidate gene approach and a concrete field application. *Annals of the New York Academy of Sciences*. 2002;969:92-6.
- [53] Juliarena MA, Poli M, Sala L, Ceriani C, Gutierrez S, Dolcini G, et al. Association of BLV infection profiles with alleles of the BoLA-DRB3.2 gene. *Animal Genetics*. 2008;39(4):432-8.
- [54] Miyasaka T, Takeshima SN, Jimba M, Matsumoto Y, Kobayashi N, Matsuhashi T, et al. Identification of bovine leukocyte antigen class II haplotypes associated with variations in bovine leukemia virus proviral load in Japanese Black cattle. *Tissue Antigens*. 2013;81(2):72-82.
- [55] Nagaoka Y, Kabeya H, Onuma M, Kasai N, Okada K, Aida Y. Ovine MHC class II DRB1 alleles associated with resistance or susceptibility to development of bovine leukemia virus-induced ovine lymphoma. *Cancer Research*. 1999;59(4):975-81.
- [56] Xu A, van Eijk MJ, Park C, Lewin HA. Polymorphism in BoLA-DRB3 exon 2 correlates with resistance to persistent lymphocytosis caused by bovine leukemia virus. *Journal of Immunology* 1993. p. 6977-85.

- [57] Schwab AE, Geary TG, Baillargeon P, Schwab AJ, Fecteau G. Association of BoLA DRB3 and DQA1 alleles with susceptibility to *Neospora caninum* and reproductive outcome in Quebec Holstein cattle. *Veterinary Parasitology*. 2009;165(1-2):136-40.
- [58] Glass EJ, Crutchley S, Jensen K. Living with the enemy or uninvited guests: functional genomics approaches to investigating host resistance or tolerance traits to a protozoan parasite, *Theileria annulata*, in cattle. *Veterinary Immunology and Immunopathology*. 2012;148(1-2):178-89.
- [59] Buitkamp J, Filmether P, Stear MJ, Epplen JT. Class I and class II major histocompatibility complex alleles are associated with faecal egg counts following natural, predominantly *Ostertagia circumcincta* infection. *Parasitology Research*. 1996;82(8):693-6.
- [60] Stear MJ, Bairden K, Bishop SC, Buitkamp J, Epplen JT, Gostomski D, et al. An ovine lymphocyte antigen is associated with reduced faecal egg counts in four-month-old lambs following natural, predominantly *Ostertagia circumcincta* infection. *International Journal for Parasitology*. 1996;26(4):423-8.
- [61] Liu LX, Zhao SG, Lu HN, Yang QL, Huang XY, Guo SB. Association between polymorphisms of the swine MHC-DQA gene and diarrhoea in three Chinese native piglets. *International Journal of Immunogenetics*. 2015;42(3):208-16.
- [62] Bao WB, Ye L, Zi C, Liu L, Zhu J, Pan ZY, et al. Relationship between the expression level of SLA-DQA and *Escherichia coli* F18 infection in piglets. *Gene*. 2012;494(1):140-4.
- [63] Goto RM, Wang Y, Taylor RL, Jr., Wakenell PS, Hosomichi K, Shiina T, et al. BG1 has a major role in MHC-linked resistance to malignant lymphoma in the chicken. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(39):16740-5.
- [64] Briles WE, Stone HA, Cole RK. Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. *Science*. 1977;195(4274):193-5.
- [65] Lillehoj HS, Ruff MD, Bacon LD, Lamont SJ, Jeffers TK. Genetic control of immunity to *Eimeria tenella*. Interaction of MHC genes and non-MHC linked genes influences levels of disease susceptibility in chickens. *Veterinary Immunology and Immunopathology*. 1989;20(2):135-48.
- [66] Banat GR, Tkalcic S, Dzielawa JA, Jackwood MW, Saggese MD, Yates L, et al. Association of the chicken MHC B haplotypes with resistance to avian coronavirus. *Developmental and Comparative Immunology*. 2013;39(4):430-7.
- [67] MacLachlan NJ, Dubovi EJ. Chapter 14: Retroviridae. In: Press A, editor. *Fenner's Veterinary Virology*. 4th ed.: Elsevier Science; 2011. p. 243-74.
- [68] USDA. Bovine Leukosis Virus (BLV) on U.S. Dairy Operations, 2007. 2008.
- [69] Samagh BS, Kellar JK, editor. Seroepidemiological survey of bovine leukaemia virus infection in Canadian cattle. *Fourth International Symposium on Bovine Leukosis*; 1982.

[70] Lorenz RJS, O.C. The epidemiology of enzootic bovine leukosis. In: Burny AaM, M., editor. *Developments in Veterinary Virology: Enzootic Bovine Leukosis and Bovine Leukemia Virus*. Boston: Martinus Nijhoff; 1987. p. 52-68.

[71] Maresca C, Costarelli S, Dettori A, Felici A, Iscaro C, Feliziani F. Enzootic bovine leukosis: report of eradication and surveillance measures in Italy over an 8-year period (2005–2012). *Preventive Veterinary Medicine*. 2015;119(3–4):222-6.

[72] Nuotio L, Rusanen H, Sihvonen L, Neuvonen E. Eradication of enzootic bovine leukosis from Finland. *Preventive Veterinary Medicine*. 2003;59(1–2):43-9.

[73] Schwartz I, Levy D. Pathobiology of bovine leukemia virus. *Veterinary Research*. 1994;25(6):521-36.

[74] Hasselschwert DL, French DD, Hribar LJ, Luther DG, Leprince DJ, Van der Maaten MJ, et al. Relative susceptibility of beef and dairy calves to infection by bovine leukemia virus via tabanid (Diptera: Tabanidae) feeding. *Journal of Medical Entomology*. 1993;30(2):472-3.

[75] Mirsky ML, Olmstead CA, Da Y, Lewin HA. The prevalence of proviral bovine leukemia virus in peripheral blood mononuclear cells at two subclinical stages of infection. *Journal of Virology*. 1996;70(4):2178-83.

[76] Panei CJ, Takeshima SN, Omori T, Nunoya T, Davis WC, Ishizaki H, et al. Estimation of bovine leukemia virus (BLV) proviral load harbored by lymphocyte subpopulations in BLV-infected cattle at the subclinical stage of enzootic bovine leucosis using BLV-CoCoMo-qPCR. *BMC Veterinary Research*. 2013;9:95.

[77] Radostitis O, Gay C, Hinchcliff K, Constable P. Chapter 21: Disease associated with viruses and chlamydia I. In: Elsevier, editor. *Veterinary Medicine*. 10th ed.: Elsevier; 2007. p. 1209-21.

[78] Hopkins SG, DiGiacomo RF. Natural transmission of bovine leukemia virus in dairy and beef cattle. *Veterinary Clinics of North America: Food Animal Practice*. 1997;13(1):107-28.

[79] Buxton BA, Schultz RD. Factors affecting the infectivity of lymphocytes from cattle with bovine leukosis virus. *Canadian Journal of Comparative Medicine*. 1984;48(4):365-9.

[80] Bartlett PC, Sordillo LM, Byrem TM, Norby B, Grooms DL, Swenson CL, et al. Options for the control of bovine leukemia virus in dairy cattle. *Journal of the American Veterinary Medical Association*. 2014;244(8):914-22.

[81] Buxton BA, Hinkle NC, Schultz RD. Role of insects in the transmission of bovine leukosis virus: potential for transmission by stable flies, horn flies, and tabanids. *American Journal of Veterinary Research*. 1985;46(1):123-6.

[82] Erskine RJ, Bartlett PC, Byrem TM, Render CL, Febvay C, Houseman JT. Herd-level determinants of bovine leukaemia virus prevalence in dairy farms. *Journal of Dairy Research*. 2012;79(4):445-50.

- [83] Kobayashi S, Hidano A, Tsutsui T, Yamamoto T, Hayama Y, Nishida T, et al. Analysis of risk factors associated with bovine leukemia virus seropositivity within dairy and beef breeding farms in Japan: a nationwide survey. *Research in Veterinary Science*. 2014;96(1):47-53.
- [84] Ooshiro M, Konnai S, Katagiri Y, Afuso M, Arakaki N, Tsuha O, et al. Horizontal transmission of bovine leukemia virus from lymphocytotic cattle, and beneficial effects of insect vector control. *Veterinary Record*. 2013;173(21):527.
- [85] Romero CH, Cruz GB, Rowe CA. Transmission of bovine leukaemia virus in milk. *Tropical Animal Health and Production*. 1983;15(4):215-8.
- [86] Dimmock CK, Chung YS, MacKenzie AR. Factors affecting the natural transmission of bovine leukaemia virus infection in Queensland dairy herds. *Australian Veterinary Journal*. 1991;68(7):230-3.
- [87] Lassauzet ML, Johnson WO, Thurmond MC, Stevens F. Protection of colostral antibodies against bovine leukemia virus infection in calves on a California dairy. *Canadian Journal of Veterinary Research*. 1989;53(4):424-30.
- [88] Ott SL, Johnson R, Wells SJ. Association between bovine-leukosis virus seroprevalence and herd-level productivity on US dairy farms. *Preventive Veterinary Medicine*. 2003;61(4):249-62.
- [89] Rodriguez SM, Florins A, Gillet N, de Brogniez A, Sanchez-Alcaraz MT, Boxus M, et al. Preventive and therapeutic strategies for bovine leukemia virus: lessons for HTLV. *Viruses*. 2011;3(7):1210-48.
- [90] Gutierrez SE, Dolcini GL, Arroyo GH, Rodriguez Dubra C, Ferrer JF, Esteban EN. Development and evaluation of a highly sensitive and specific blocking enzyme-linked immunosorbent assay and polymerase chain reaction assay for diagnosis of bovine leukemia virus infection in cattle. *American Journal of Veterinary Research*. 2001;62(10):1571-7.
- [91] Juliarena M, Gutierrez S, Ceriani C. Chicken antibodies: a useful tool for antigen capture ELISA to detect bovine leukaemia virus without cross-reaction with other mammalian antibodies. *Veterinary Research Communications*. 2007;31(1):43-51.
- [92] Juliarena MA, Gutierrez SE, Ceriani C. Determination of proviral load in bovine leukemia virus-infected cattle with and without lymphocytosis. *American Journal of Veterinary Research*. 2007;68(11):1220-5.
- [93] Dube S, Abbott L, Dube DK, Dolcini G, Gutierrez S, Ceriani C, et al. The complete genomic sequence of an in vivo low replicating BLV strain. *Virology Journal*. 2009;6:120.
- [94] Burny A, Cleuter Y, Kettmann R, Mammerickx M, Marbaix G, Portetelle D, et al. Bovine leukaemia: facts and hypotheses derived from the study of an infectious cancer. *Cancer Surveys*. 1987;6(1):139-59.
- [95] Juliarena MA, Lendez PA, Gutierrez SE, Forletti A, Rensetti DE, Ceriani MC. Partial molecular characterization of different proviral strains of bovine leukemia virus. *Archives of Virology*. 2013;158(1):63-70.

- [96] Croshaw JE, Jr., Abt DA, Marshak RR, Hare WC, Switzer J, Ipsen J, et al. Pedigree studies in bovine lymphosarcoma. *Annals of the New York Academy of Sciences*. 1963;108:1193-202.
- [97] Ferrer JF. Bovine lymphosarcoma. *Advances in Veterinary Science and Comparative Medicine*. 1980;24:1-68.
- [98] Lewin HA, Bernoco D. Evidence for BoLA-linked resistance and susceptibility to subclinical progression of bovine leukaemia virus infection. *Animal Genetics*. 1986;17(3):197-207.
- [99] Palmer C, Thurmond M, Picano J, Brewer AW, Bernoco D, editor. Susceptibility of cattle to bovine leukemia virus infection associated with BoLA type. 91st Annual Meeting of the United States Animal Health Association; 1987; Richmond, VA, EUA: Animal Health Association.
- [100] Lewin HA, Wu MC, Stewart JA, Nolan TJ. Association between BoLA and subclinical bovine leukemia virus infection in a herd of Holstein-Friesian cows. *Immunogenetics*. 1988;27(5):338-44.
- [101] Stear MJ, Dimmock CK, Newman MJ, Nicholas FW. BoLA antigens are associated with increased frequency of persistent lymphocytosis in bovine leukaemia virus infected cattle and with increased incidence of antibodies to bovine leukaemia virus. *Animal Genetics*. 1988;19(2):151-8.
- [102] Lewin HA. Disease resistance and immune response genes in cattle: strategies for their detection and evidence of their existence. *Journal of Dairy Science*. 1989;72(5):1334-48.
- [103] Zanotti M, Poli G, Ponti W, Polli M, Rocchi M, Bolzani E, et al. Association of BoLA class II haplotypes with subclinical progression of bovine leukaemia virus infection in Holstein-Friesian cattle. *Animal Genetics*. 1996;27(5):337-41.
- [104] Zavala-Ruiz Z, Sundberg EJ, Stone JD, DeOliveira DB, Chan IC, Svendsen J, et al. Exploration of the P6/P7 region of the peptide-binding site of the human class II major histocompatibility complex protein HLA-DR1. *The Journal of Biological Chemistry*. 2003;278(45):44904-12.
- [105] Tiwari JL, Terasaki PI. HLA-DR and disease associations. *Progress in Clinical and Biological Research*. 1981;58:151-63.
- [106] Aida Y. Influence of host genetic differences on leukemogenesis Induced by bovine leukaemia virus. *AIDS Research and Human Retroviruses* 2001; 17 (supplement 1), S12. doi:10.1089/088922201750252223.
- [107] Mirsky ML, Olmstead C, Da Y, Lewin HA. Reduced bovine leukaemia virus proviral load in genetically resistant cattle. *Animal Genetics*. 1998;29(4):245-52.
- [108] Juliarena M. Contribution to the study of the role of virus and host in the infection profile caused by Bovine Leukemia Virus. Doctoral Thesis FCV-UNCPBA. 2008.

[109] Esteban EN, Poli M, Poiesz B, Ceriani C, Dube S, Gutierrez S, et al. Bovine leukemia virus (BLV), proposed control and eradication programs by marker assisted breeding of genetically resistant cattle. In: Rechi LJ, editor. *Animal Genetics*. Hauppauge, NY, USA: Nova Science Publishers, Inc.; 2009. p. 107-30.

[110] Juliarena MA, Poli M, Ceriani C, Sala L, Rodriguez E, Gutierrez S, et al. Antibody response against three widespread bovine viruses is not impaired in Holstein cattle carrying bovine leukocyte antigen DRB3.2 alleles associated with bovine leukemia virus resistance. *Journal of Dairy Science*. 2009;92(1):375-81.

[111] Rhodes JK, Pelzer KD, Johnson YJ. Economic implications of bovine leukemia virus infection in mid-Atlantic dairy herds. *Journal of the American Veterinary Medical Association*. 2003;223(3):346-52.

[112] Esteban EN, Thorn RM, Ferrer JF. Characterization of the blood lymphocyte population in cattle infected with the bovine leukemia virus. *Cancer Research*. 1985;45(7):3225-30.

[113] Bendixen HJ. Bovine enzootic leukosis. *Advances in Veterinary Science* 1965;10:129-204.

[114] Revelli GRS, O.A. Prevalencia de leucosis enzoótica bovina en la zona noroeste de Santa Fe y sur de Santiago del Estero, Argentina. *Revue de Médecine Vétérinaire (Bogotá)*. 2004;85(4):135-8.

[115] Barrios CN, Rensetti DE, Revelli GR, Ceriani MC, Trabattoni E, Esteban EN, Juliarena M, editor. En 2012, al norte del paralelo 34°S, el virus de la leucosis bovina infectaría más del 99% de los tambos con endemias en rodeos de carne. Reunión Anual de la Sociedad Argentina de Virología; 2012.

[116] Juliarena MA, Barrios CN, Ceriani MC, Esteban EN. Hot topic: Bovine leukemia virus (BLV)-infected cows with low proviral load are not a source of infection for BLV-free cattle. *Journal of Dairy Science*. 2016;99(6):4586-9.

[117] Juliarena MA, Rensetti D, Favier P, Morán PE, Pérez SE, Carignano HA, Basigalupo D, Poli M, Esteban EN, Gogorza LM. Study of the association between the presence of BoLA DRB3\*0902 and infection by bovine viral diarrhea virus, bovine herpesvirus-4 and neosporacaninum. *Revista Argentina de Microbiología*. 2011;Supl 1:140.

[118] Lützelschwab CM, Forletti A, Cepeda R, Esteban EN, Confalonieri O, Gutiérrez SE. Co-infection with *Mycobacterium bovis* does not alter the response to bovine leukemia virus in BoLA DRB3\*0902, genetically resistant cattle. *Research in Veterinary Science*. <http://dx.doi.org/10.1016/j.rvsc.2016.09.002>

[119] Dietz AB, Detilleux JC, Freeman AE, Kelley DH, Stabel JR, Kehrli ME, Jr. Genetic association of bovine lymphocyte antigen DRB3 alleles with immunological traits of Holstein cattle. *Journal of Dairy Science*. 1997;80(2):400-5.

[120] Rupp R, Hernandez A, Mallard BA. Association of bovine leukocyte antigen (BoLA) DRB3.2 with immune response, mastitis, and production and type traits in Canadian Holsteins. *Journal of Dairy Science*. 2007;90(2):1029-38.

- [121] Kelm SC, Detilleux JC, Freeman AE, Kehrli ME, Jr., Dietz AB, Fox LK, et al. Genetic association between parameters of innate immunity and measures of mastitis in periparturient Holstein cattle. *Journal of Dairy Science*. 1997;80(8):1767-75.
- [122] Bojarcz-Nosowicz BK, E. Somatic cell count and chemical composition of milk in naturally BLV-infected cows with different phenotypes of blood leukocyte acid phosphatase. *Archiv Tierzucht, Dummerstorf*. 2006;49(1):17-28.
- [123] Abdalla EA, Weigel KA, Byrem TM, Rosa GJM. Short communication: Genetic correlation of bovine leukosis incidence with somatic cell score and milk yield in a US Holstein population. *Journal of Dairy Science*. 2016;99(3):2005-9.
- [124] Yoshida T, Furuta H, Kondo Y, Mukoyama H. Association of BoLA-DRB3 alleles with mastitis resistance and susceptibility in Japanese Holstein cows. *Animal Science Journal*. 2012;83(5):359-66.

---

## New Advances in NGS Technologies

---

Edo D'Agaro

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/66924>

---

### **Abstract**

In the next-generation sequencing (NGS) methods, a DNA molecule of an individual is broken down into many small fragments to make up the so-called sequencing library. These small fragments serve as a template for the synthesis of numerous complementary fragments (called reads). Every small piece of the original DNA is copied many times in a variable number of reads. Depending on the desired accuracy level, it is possible to set the system to achieve a certain level of coverage, i.e., a number of reads per fragment. A level of 30X coverage is already sufficient for the routine diagnosis of most of the Mendelian diseases. All the sequences are then transferred into a computer and aligned with a reference sequence available in the international databases. By this way, all sequences of reads can be recomposed as a fine puzzle to obtain the sequence of a single gene or whole genome. The NGS machines, available today, are very flexible devices. In fact, an NGS sequencer can be used for different types of applications: (1) whole-genome sequencing (WGS): analysis of the entire genome of an individual; (2) whole exome sequencing (WES): analysis of the entire coding genes of an individual; (3) targeted sequencing: analysis of a set of genes or a single gene; (4) transcriptome analysis: analysis of all the RNA produced by specific cells.

**Keywords:** genomics, NGS, genetic screening, animal diseases, bioinformatics

---

### **1. Introduction**

The next-generation sequencing (NGS) technologies, with the capability to produce, at a low price, millions of DNA sequences per analysis, have changed the way to make the genetic analysis in the veterinary and animal fields. In the past few years, the observed reduction of cost and time required to obtain the genomic information (whole genome, whole exome or targeted gene sequencing) have had several repercussions in various disciplines such as animal selection, animal research, genetic disease diagnosis and control, cancer research, and metagenomic studies. The focus of this chapter is to give an updated overview of the next-generation

techniques, analyzing the general framework, the bioinformatic analysis pipeline, present applications in the genetic disease detection, problems and future possible developments.

## 2. Patterns of transmission of genetic diseases

### 2.1. The hereditary transmission

The hereditary traits are transmitted from the DNA (deoxyribonucleic acid) contained in the chromosomes and in the mitochondria. The gene is formed by regions that encode proteins, the exons, and segments that are not translated, the introns. Genetic diseases are due to inherited mutations, which are changes in DNA sequence and transmitted in familiars in different ways. Several diseases are inherited in a simple way and follow the Mendel's law. These are monogenic or Mendelian diseases. Monogenic diseases can be classified according to their mechanisms of inheritance: autosomal dominant and recessive, X-linked and mitochondrial diseases.

### 2.2. Autosomal dominant

In the case of dominant disease, the most common genotype is the heterozygous genotype. The disease is transmitted to an average of 50% of the offspring regardless of their sex. The most common cross is between a heterozygous affected (Aa) and a homozygous unaffected (aa). A typical pedigree of an autosomal dominant disease has the following distinguish features:

- (a) vertical segregation, with affected individuals in all generations, both males and females;
- (b) an equal number of affected males and females;
- (c) each affected individual have at least one affected parent.

An autosomal dominant disease may also arise from a new mutation. A risk factor of a new mutation is the parental age. The frequency of new mutations increases proportionally with the age of the parents. One feature of the dominant autosomal diseases is to show a variable expression among the affected individuals, even in the same family. Some autosomal dominant diseases exhibit the phenomenon of anticipation, which consists of an earlier onset and an aggravation of symptoms with the passing generations. Another characteristic of the autosomal dominant diseases is the penetrance, which is the percentage of animals expressing the phenotype over the total number of individuals which carry the mutation. Other genes or the environment can interfere with an occurrence of a disease resulting in a lack of penetrance. In this case, the subject that carries the dominance mutation will not be affected but, 50% of the offspring will manifest the disease with varying degrees.

### 2.3. Autosomal recessive diseases

The presence of two recessive alleles at the same locus causes the disorders. The most frequent cross is between heterozygous individuals for the mutated allele (healthy carriers of an autosomal recessive disease). A typical pedigree of an autosomal recessive disease has the following distinguish features:

- (a) unaffected parents could have affected offspring;
- (b) equal number of males and females;
- (c) all offspring are affected when parents are affected.

The heterozygous parents will show a risk of 25% to get affected offspring. Inbreeding is the most important risk factor for the occurrence of an autosomal recessive disease.

#### **2.4. Disease linked to the chromosome X**

The X chromosome contains hundreds of genes. The dominant or recessive effect is determined in the females because they have two chromosomes. Males have only one X chromosome and therefore they express the recessive mutations (called hemizygous). In this case, the recessive trait is always expressed. If the alleles are different, a female is heterozygous, i.e., a carrier. A typical pedigree of an x-linked dominant disease has the following distinguish features:

- (a) each generation usually has an affected individual;
- (b) all daughters of an affected male are affected;
- (c) all males and females of a heterozygous are affected.

The family tree study shows a zig-zag or diagonal pattern, with a variable number of affected males in the following generations.

A typical pedigree of an x-linked recessive disease has the following distinguish features:

- (a) all males of an affected female are affected;
- (b) affected males never transmit the trait to their male offspring;
- (c) unaffected parents may have affected offspring.

#### **2.5. Multifactorial diseases**

Some genetic diseases show a polygenic and/or multifactorial inheritance. The multifactorial traits can be continuous and measurable, such as the weight, height, blood pressure, or discontinuous, such as congenital heart disease, diabetes, osteoporosis, and cardiovascular disease.

#### **2.6. Mitochondrial disease**

The defects in the mitochondrial DNA show a maternal inheritance because mitochondria are transmitted to the offspring by the egg cell. Characteristics of mitochondrial diseases are incomplete penetrance, variable expressivity, and pleiotropy. The phenotypic expression of mitochondrial disorders depends on the ratio of normal and mutated mtDNA present in the cells of various tissues. The organ impairment and its severity depend on the percentage of mutated mitochondria, which is variable in different tissues. Mitochondria are distributed in cells throughout the body and are therefore responsible for pathologies with impairment of different organs and apparatuses.

### 3. Mutations

A mutation is defined as a stable and heritable modification of the genetic material. The mutation can occur at different levels and change the attitude of the entire genome showing a variation in the number of chromosomes (genomic mutation), a variation in a single chromosome (mutation chromosome) or, it may involve a single gene (gene mutation). Mutations can be spontaneous or induced. The most common causes of mutation can be physical, chemical, or biological factors (called mutagens). The gene mutation can be lethal to the cell or, when the gene is expressed, give rise to senseless, inactive, or less active proteins or, although much rarely, resulted in products with a greater or different activity. Most of the mutations are harmful and only a small percentage appears to be advantageous. The carrier of an unfavorable gene may die before to reproduce (lethal disease). Natural selection tends to reduce mutated genes. Sometimes, however, due to the changing environment, insignificant mutations can be expressed or even turn out to be advantageous and to be favored by natural selection. Genomic and chromosomal mutations, when they are compatible with life, give rise to a more complex phenotypic pattern, because they involve a large number of genes. On the contrary, gene mutations affect the sequence of a single gene and therefore they have an effect only on the functionality of a single protein. A *transition mutation* is one substitution of a purine with a purine or pyrimidine with a pyrimidine while a *transversion mutation* is a substitution between a purine and a pyrimidine. At phenotypic level, effects of mutations can vary from the complete neutrality to a range of events whose gravity depends on the interaction with other genes and with the environment. At the molecular level, a gene mutation consists in a variation of the sequence of a nucleotide of the gene. The normal DNA sequence may change as a result of insertion or deletion of one or more bases. A gene mutation, if it concerns a single point of DNA, i.e., one or a few bases, it is called a point mutation or SNP. The first element to be considered, in assessing the effects of a gene mutation, is what part of the gene has changed. The sequence of a gene is made, in addition to a promoter, of a number of exons interspersed with introns. As is known, the entire sequence of the gene is transcribed, but the intronic fragments are removed during the process of splicing. The exon region of a gene is characterized by the presence of a coding region, which contains the information for the synthesis of the coded protein, and two noncoding regions that are located at the beginning and end of the mRNA molecule (5'UTR and 3'UTR that are transcribed but not translated). Mutations that modify the coding region may have important effects on the translation of the proteins, mutations at the junction between intron and exon can interfere with the recognition of splice sites causing the production of modified transcripts while mutations in the untranslated regions (UTRs) and the introns are often neutral or in some cases may alter the regulatory functions of these regions. The mutations that show (most frequently) phenotypic effects are those affecting the coding regions. Therefore, in this chapter, we will focus on them. In the case of a *silent mutation*, the replacement of a base causes a change in the codon without a change of the amino acid (the genetic code is degenerated and the same amino acid can be coded by more than one codon). For example, if a mutation changes the codon CUA in CUG, as both codons code for the same amino acid leucine, the amino acid does not change. For this reason, a silent mutation has no effect at the level of the protein and thus on the phenotype. In the case of a *missense mutation*, the substitution of a base of a codon with a

different base causes the change in the codon function. For example, if the codon AGC, which codes for the amino acid serine, becomes AGA, encoding for the amino acid arginine, during the protein synthesis, an arginine will be inserted instead of a serine. The substitution of an amino acid may therefore change the protein function. Based on the amino acid substitution, the missense mutations are divided into: (a) conservative substitutions: the new amino acid has similar characteristics to the replaced one (neutral mutation); (b) nonconservative substitutions: the new amino acid has different characteristics [e.g. GAG (Glu) changing to GTG (Val)]. The consequences of a missense mutation are more or less serious depending on the physical-chemical characteristics of the two amino acids involved. Characteristics of the amino acids present in the primary structure determine the secondary and tertiary structures of the proteins by means of ionic and hydrophobic interactions. In turn, the correct folding of the protein chain determines its biological functionality. The maintenance of the protein functions depends on the amino acid position and its structural role. In the case of sickle cell disease, due to a missense mutation in the gene of the beta-globin, a polar amino acid (water-soluble, glutamic acid) is replaced by an insoluble apolar amino acid valine. This causes the production of a beta-globin chain resulting in an altered hemoglobin solubility, which in turn causes the characteristic shape of the sickle red blood cells. A gene mutation can change a codon sense in a noncodon sense (*nonsense mutation*). For example, if the codon AAG, coding for lysine, becomes UAG, which is a stop codon, then, the protein synthesis will end prematurely. The synthesized protein is therefore incomplete and, in most cases, it will not be functional. An elongation mutation arises when, a nonsense triplet present in the wild gene, following a base substitution, becomes a triplet "sense," encoding for an amino acid. In this case, the stop triplet is eliminated and the mRNA translation will continue with the formation of a longer protein. Deletions and insertions of bases in the coding region of a gene cause a slippage of the reading code system (*frameshift mutations*). The inclusion or loss of one or more bases causes a slip, resulting in loss of sense of the whole protein. All codons downstream of the mutation change and from that point they are then incorporated into the protein, giving rise to an abnormal protein and almost always nonfunctional. A deletion of an entire portion of a gene can severely alter the three-dimensional protein structure, the chemical properties and functionality. Even in this case, the deletion may lead to a slip of reading (*frameshift*), determining the loss of a part of the protein. A mutation in a gene can also determine a quantitative variation (not qualitative) of the protein. This occurs when the mutation falls in a regulatory region (i.e., the sequence of the promoter) and modifies the transcription of the gene. A point mutation (by substitution or by insertion/deletion of bases), both in coding and noncoding regions, can give rise to an RFLP (polymorphism of length of the restriction fragments) and to an SNP (single nucleotide polymorphism).

## 4. DNA sequencing

Over the last 60 years, it has observed a significant increase in the knowledge of animal genomes and the genetic code, starting from the discovery of the structure of DNA in 1953 until the publication of the first draft of the human genome in 2001 [1, 2]. The Sanger sequencing [3], also known as sequencing of the first generation, was the method used to sequence

the genome within the project "Human Genome" bringing the entire genome sequence in 2003, after 13 years from the beginning of the project with a cost of \$3 billion and the contribution of six different nations. In the recent years, several animal and plant genome sequencing projects have been carried out thanks to the implementation of both molecular biology and computer science innovations [4]. The evolution in parallel of these two sectors has enabled the advent of several next-generation sequencing (NGS) platforms. Nowadays, by means of this new technology, which is more efficient and economic compared to the previous methods, it is possible to sequence a new genome at the cost of less than one thousand euros in a very short time (1–2 days). Furthermore, the cost and time can be greatly reduced by analyzing only the coding regions (exomes). Despite the fact that the exomes are only 1% of the entire genome, by means of this method, it is possible to identify the 85% of the monogenic diseases.

#### 4.1. Last-generation sequencing (NGS) possible

All NGS methods are characterized by two important phases: a molecular biology step, which goes from the preparation of the sample to the sequencing phase, which allows to carry out, at the same time, more than one reaction and with less dexterity than the Sanger method, and a computerized stage for the analysis of the data. The first part of the process is divided into three steps: *sample preparation*, *amplification*, and *sequencing*. Sample preparation, which is the common passage for all platforms, is the fragmentation of DNA (dimensions vary from 100 to 800 base pairs (bp) in relation to the platform and the adapters used). The amplification is based on two methods: emulsion or solid PCR amplifications. In the emulsion PCR method, described for the first time by Tawfik Griffiths, the individual molecules of DNA are clonally amplified in micro-compartments consisting of mixtures of water and oil [5]. The adapters bound to individual DNA molecules, hybridize to complementary sequences coating the surface. Cycles of upcoming amplification allow to the formation of "clusters" of folded fragments clonally amplified. All the NGS platforms are characterized by the ability to sequence and massively parallel amplifying the DNA molecules in a clonal or single way. Unlike the Sanger method, where the fragments of different sizes from individual sequencing reactions were separated by electrophoresis, in the NGS technologies, sequencing is accomplished through the repetition of nucleotide extension cycles or oligonucleotide ligations. The principles of sequencing and image acquisition are the peculiar steps that characterize the different platforms on the market. To date, different methods of sequencing are known. The cyclic reversible termination (CRT) sequencing method uses reversible labeled nucleotides. Each sequencing cycle comprises: the incorporation of the nucleotide, the acquisition of the fluorescence, and the cutting of the nucleotide. The subsequent washing step allows the elimination of all nonincorporated nucleotides. At this point, the image for identifying the incorporated nucleotide is captured, followed by a cleavage step that removes the terminator group and the fluorophore. After the elimination, the polymerase can continue the reaction of extension and tie the second nucleotide. This process is used by two types of trading platforms: Illumina and Helicos, which differ in the templates of sequencing. While Illumina uses clonally amplified fragments on a solid surface, Helicos is currently the only commercial platform able to use single DNA molecules. In addition, the acquisition platform of Illumina uses a four-color (the four reversible nucleotides are labeled with a different fluorophore and

are dispensed at the same time in the sequencer), while in the Helicos platform, all nucleotides they are labeled with the same fluorophore and are dispensed into the sequencer in a determined hierarchical order.

#### 4.2. Commercial systems for NGS

We describe here the main platforms used for the last-generation sequencing:

*Roche 454 system* was the first genome sequencer to be marketed in 2004. This system uses the sequencing by synthesis technology known as pyrosequencing. Initially, the 454 sequencing method used reads of 100–150 bp, producing about 200,000 reads, with an output of 20 Mb for each run. In 2008, a new sequencer the 454 GS FLX titanium was produced (reads of 700 bp, with an accuracy after filtering of 99.9%, with an output of 0.7 Gb for run in 24 hours). In 2009, the GS Junior and 454 GS FLX platforms were used to set up a new sequencer with an output of 14 Gb for run. Further developments have led to the production of the GS FLX +, which is able to sequence reads up to 1 kb. The high speed of analysis combined with the long reads is the positive characteristics of this platform. However, the cost of reagents remains a problem to be solved.

*Ion Torrent*. The method used by Ion Torrent Genome Machine (PGM) is very similar to the Roche 454 system. This platform, instead of using images to capture the incorporated nucleotide, detects the change of pH. The output per run is of 270 Mb with reads of 100–200 bp.

*AB SOLiD system*. The Applied Biosystems method was first marketed in 2006. This system uses the ligation sequencing method (in both directions) to ligate fluorescently labeled octomers to the DNA fragment. Initially, the length of the reads was only 35 bp and the output of 3 Gb for run. In 2010, the 5500x1 SOLiD platform was released (reads with length of 85 bp, precision of 99.99% and output 30 Gb per run). The main problem of this method is the low length of the reads.

*Illumina GA/HiSeq system*. In 2006, the Solexa company released the Genomic Analyzer (GA) and in 2007, the company was bought by Illumina. The system uses the sequencing by synthesis method (SBS) and the amplification bridge which is an alternative method of the PCR. At the beginning, the analyzer's output was of 1 Gb for run, after upgraded to 50 Gb. In 2010, the HiSeq 2000 was released with an output of 600 Gb for run and reads of 200 bp. The basic method is to generate a great number of colonies (generated DNA colonies) that are simultaneously sequenced in parallel reactions occurring on the surface of a flow cell. The cost of sequencing is lower compared to the competitors. **Table 1** presents the main NGS platforms.

##### 4.2.1. Third-generation sequencing technologies.

*Helicos single-molecule sequencing device*: Heliscope method was presented for the first time in 2007. This platform uses a technique that individually analyzes the molecules achieving a greater accuracy. Using this system, you can obtain an output in the order of 28 Gb. However, the main disadvantage of this method remains the low capacity to correctly identify indels,

Platform	Type
Illumina	Sequencing by synthesis
454 Roche	Pyrosequencing
Ion Torrent	Ion semiconductor
ABI Solid	Ligation based sequencing (color spaced)
Oxford nanopore	Nanopore sequencing
Pacific biosciences	Single molecule real time (SMRT)

**Table 1.** List of sequencing platforms.

with a consequent increase in errors. Another problem is the length of the reads, which has never exceeded 50 bp.

*Oxford Nanopore Miniontechnology.* Most of the nanopore sequencing technologies are based on the transit of DNA molecules, after the application of an electric potential, through an array of protein nanopores located in a high-salt buffer. Single bases are identified according to the electric change. Several methods have been proposed based on the nanopore technique. Among these, one is produced by Oxford Nanopore (exploits the combination of three molecules), one based on MSpa (a protein). This method can be very useful for identifying infectious diseases and to analyze environmental samples (metagenomic analysis of bacteria, viruses and fungi) without a previous knowledge of the sample composition.

An additional advantage of the third-generation sequencing methods is the production of long reads (>than 10,000 pb and up to 100,000 bp) enabling to analyze insertions, deletions, and translocations [6–9]. Using these technologies and mapping methods, it is possible to study entire chromosome arms.

### 4.3. NGS data analysis

Although the various sequencing technologies use different methods, they all provide as output the FASTQ sequences. The FASTQ string is the type of information used in molecular biology to store genetic sequences and the related quality scores, namely the score that the algorithm assigns to the string and which is then used to choose the best match against the reference genome. At this point, the bioinformatics analysis is divided into three steps: *alignment*, which search of correspondences between the reads and the reference genome and the *variant calling* that attempts to separate the differences due to genetic mutations and instrumental errors made during the analysis and *filtration and annotation*, which attempt to align the reads to the reference genome.

#### 4.3.1. Alignment

The alignment is the process by which you map short reads to a reference genome. It is a complex task, since the software must compare each reads in all of the reference DNA positions

[8, 10, 11]. It is a computationally challenging passage and wasteful in terms of time. The SAM (sequence alignment map) and BAM (binary alignment map) are the standard file formats for storing the data obtained using the next-generation technologies (NGS). There are many commercially available software, free or on sale, to perform this task. Most software use a method based on indices, which are very fast in the search for all alignment positions (without gaps) in the reference genome. Other algorithms, instead, allow the search for alignments with gaps. The various methods to solve the problem include the use of hash tables (e.g., MAQ, ELAND), algorithms based on Burrows-Wheeler transformation (e.g., BWA, Bowtie, SOAP2), an algorithm that uses a reversible compression method commuting the order of the characters without changing the values; genome-based hash (e.g., Novoalign, SOAP). Some software can take into account of gaps (e.g., BWA, Bowtie2) while others do not (e.g., MAQ, Bowtie).

#### 4.3.2. *Variant calling*

After the alignment, the DNA sequence can be compared to the reference genome, identifying the possible changes. These variations can be due to genetic diseases or they can be only noise. The complexity of this subject lies in the difficulty to distinguish between the true variations and the sequencing errors [12]. The continuous development and improvement of NGS technologies has brought continuous advantages in this area, improving the quality of analysis [13, 14]. The main difficulty in this kind of analysis is the presence of indels, i.e., phenomena of inserts (insertion) or cancelation (deletion) of DNA segments. In fact, indels are the major cause of false-positives. The number of false-positives increases when using algorithms that do not take it into account of gaps. Another source of errors arises during the preparation of sequences and the PCR analysis. In order to reduce the errors, it is advisable to increase the sensitivity of the PCR analysis, use updated alignment software and a big reference database for getting wider comparisons.

#### 4.3.3. *Filtering and annotation*

After the alignment and variant calling steps, a list of thousands of potential differences is generated between the genome under study and the reference genome. The next step is then to determine which of these variations are due to sequencing errors. The use of specific filters allows to remove variants that do not follow the models under study and to make annotations. The comparison of reads against a genetic reference tree will find all elements of known function. In addition to the filtration step, the annotation process provides another tool to select and restrict the test sample, applying specific functional models [15]. The low cost of these new instruments is leading to the discovery of a great number of genetic variants and identifying various diseases [16]. For example, it has been observed that about 1300 positions of a gene or a sequence within a chromosome are associated with 200 diseases [17].

#### 4.3.4. *Definitions of DNA sequence variants*

In general, there are three principal categories of variants:

*(a) Causative variants*

Variants showing a clear pathogenetic role and associated with a phenotype disease.

*(b) Unrelated variants*

The most commonly used definition to define this class of variants is “incidental findings.”

*(c) Variants with undefined functional and clinical effects (variants of uncertain significance, VUS).*

**Table 2** reports some useful Internet sites for the genetic variant discovery analysis.

SAMTools	<a href="http://mamtools.sourceforge.net">http://mamtools.sourceforge.net</a>
	<a href="http://htslib.org">http://htslib.org</a>
GATK	<a href="https://www.broadinstitute.org/gatk">https://www.broadinstitute.org/gatk</a>
Platypus	<a href="http://www.well.eox.ac/platypus">http://www.well.eox.ac/platypus</a>
Freebayes	<a href="http://github.com/ekg/freebayes">http://github.com/ekg/freebayes</a>
BreakDancer	<a href="http://breakdancer.sourceforge.net">http://breakdancer.sourceforge.net</a>
Dindel	<a href="https://sanger.ac.uk/resources/software/dindel">https://sanger.ac.uk/resources/software/dindel</a>

**Table 2.** List of variant discovery resources (Internet sites).

#### 4.3.5. Example of a variant calling analysis using the galaxy platform ([www.galaxy.org](http://www.galaxy.org))

The protocol in a nutshell:

- Inspect the reads;
- Raw data cleanup/quality trimming;
- Align reads to a reference genome;
- Mark duplicates;
- Combine samples into a single file;
- Realign reads around insertions and deletions (indels);
- Correct inaccurate base qualities (optional);
- Make variant calls, SNVs, indels;
- Annotate the results.

*(a) Read quality control*

Steps involved and suggested tools:

NGS: QC and manipulation → FastQC

Command line: `fastqc`

Some of the important outputs of FastQC for our purposes are:

- Quality encoding type: important for quality trimming software;
- Total number of reads: gives you an idea of coverage;
- Presence of highly recurring k-mers;

*(b) Quality trimming/cleanup of read files*

NGS: QC and manipulation → trimmomatic

Suggested trimmomatic functions:

- Adapter trimming;
- Sliding window trimming;
- Trailing bases quality trimming;
- Leading bases quality trimming;
- Minimum read length.

*(c) Genome alignment*

Suggested tools:

NGS: Mapping → BWA

NGS: Picard → MarkDuplicates

NGS: SAM Tools → Merge BAM Files

NGS: GATK Tools → Indel Realigner

NGS: GATK Tools → Count Covariates

NGS: GATK Tools → Table Recalibration

*(d) Variant calling*

NGS: GATK Tools → Unified Genotyper

Possible alternative software:

VarScan <http://varscan.sourceforge.net/>

*(e) Annotation*

NGS: GATK Tools → Variant Annotation

**Figure 1** shows an example of a variant analysis workflow.

## Variant analysis workflow

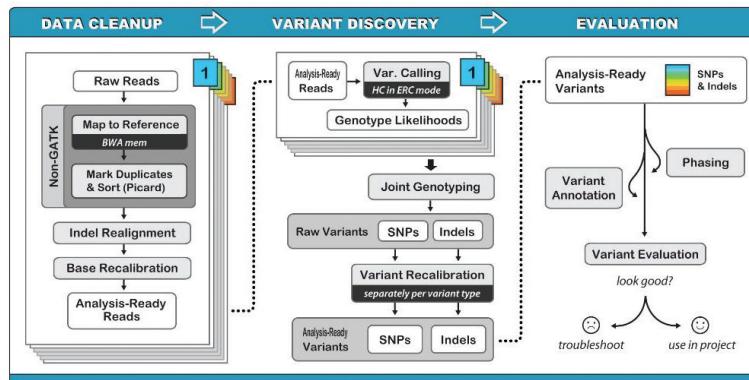



Figure 1. Example of a variant analysis workflow.

### 4.3.6. Bioinformatic analysis of the NGS data

Genetic testing based on NGS analysis can be divided depending on the width of the portion of the analyzed genome:

- sequencing the entire genome (whole genome sequencing; WGS);
- sequencing the entire exome (whole exome sequencing; WES) [18];
- targeted gene sequencing.

The choice of the different method will take into account several factors:

- *Costs*: since the introduction of the NGS practice, the use of specific panels has been privileged for economic reasons. Nowadays, the cost of the different methods has gradually narrowed and today this aspect is not, in most of the cases, the primary choice element. WES may be, in fact, less expensive of sequencing a panel of genes and, similarly, the cost of WGS is getting closer to that of WES;
- *Purpose*: in general, the NGS techniques find the best applications in the case of high heterogeneity genetic diseases or for the Mendelian diseases in which the gene is not yet been identified;
- *Diagnosis*: direct clinical applications are more and more frequent, especially for the congenital diseases;
- *Sensitivity*: the higher the number of reads of a specific region, the greater will be the sensibility for that particular stretch of DNA. In principle, in the analysis of specific panels, the smaller is the portion of the genome analyzed, the greater is the coverage and then the

sensitivity. This is the reason why, for a mosaic alteration, the analysis of known genes shows a greater sensitivity than the WES analysis;

- *Data storage*: in relation to the increasing number of sequenced genomes and data to be analyzed, new bioinformatic platforms for storage and data analysis are needed.

The amount of data generated by NGS platforms is on the order of terabytes (Tb). Software used for data analysis differ according to the NGS technology but, all follow a "pipeline" system of data analysis converting the luminescence images or fluorescence data to "reads" sequences. In this process, defined, "base calling," a quality score value is set for each nucleotide, indicating the probability of an error associated with it. The "quality score" is an important value for selecting the reads during the analysis process. If they fall below a specific level, they are eliminated improving the accuracy of the alignment process. In order to get an adequate efficiency of the alignment, this value should not be below 30. Another limitation for the alignment process is represented by the repeated sequences. The error rates associated with the NGS technologies seem to be higher than the traditional Sanger method. However, the accuracy of the NGS sequencing is increased by means of a repeated and massive reading of each gene fragment, which determines the "coverage" of the genome. The latter parameter represents an essential value in the NGS analysis (value ranging between 20 and 50 times, in relation to the platform used) determining the presence of false negatives (for heterozygous individuals) in the detection of nucleotide variants. At the end of the variant annotation process, the number of identified variants varies in relation to the application of various filters, used to reduce the number of candidates. The filters most commonly used are based on:

- inheritance pattern (autosomal/X-linked, dominant/recessive);
- sharing equal variations in well-characterized families;
- removal of variants already known through the use of public databases (dbSNP);
- based on the potential effect of the change (no sense, missense, changes in splicing sites or insertions and deletions that modify the reading frame);
- prediction of the functional effects of changes through the use of *in silico* bioinformatic tools (SIFT, POLYPHEN, and ANNOVAR software).

#### 4.3.6.1. Advantages of using the NGS technology

Although, in previous decades, the development of the Sanger method has brought many findings in chemistry, automation and miniaturization of the process, only the advent of NGS technologies has led to a significant improvement of quality and speed of the genome analysis [19]. The release of new sequencing platforms and the reduction of costs associated with the NGS technology is the consequence of three factors:

- (1) Thousands/millions of sequencing reactions can be conducted in parallel exceeding the limit of 1–96 possible reactions obtainable with a traditional sequencer;
- (2) Cloning or amplification of the DNA fragment is, in the new technologies, not necessary or completely automated within the platforms;

(3) Ability to detect the minor allele with high accuracy. This in term allows a better identification of a variant in mosaic sample or heterozygous deletions. The number of times that a DNA fragment is amplified and sequenced is proportional to the abundance of such segment in the original sample.

#### 4.3.6.2. *Limits of NGS technology*

The NGS technology has, however, some limitations related mostly to the magnitude of data products. In fact, in the NGS results we can observe false positives and false negatives.

False positives can result from:

- an incorrect alignment with the reference sequence. You can overcome this problem using different alignment software;
- systematic sequencing errors. This type of error can be identified in all the samples and removed from the final list [20];
- technical errors of the sequencer. For example, with the pyrosequencing method, a common error was observed using homopolymers longer using the pyrosequencing method, was observed when reading homopolymers longer than 5–6 bases;

False negatives instead come from:

- low coverage;
- low coverage in regions of interest;
- alignment of repeated regions.

The reduction of errors can be obtained increasing the coverage and the quality of DNA fragmentation (fragments of greater size). The implementation of the “paired-end” sequencing, that is able to sequence fragments of a greater length by both ends, allows the analysis of fragments of 5–10 kb. The use of the Sanger method is, however, required at the end of the analysis because the results, obtained using the NGS methods, require to be validated.

#### 4.3.7. *Applications of NGS technology*

The genome can be assessed globally, only in the coding regions or in target regions. The main interest of the veterinary genetic sector is to study the coding regions as the greater number of diseases is caused by mutations or exonic splicing which alter the correct amino acid sequence of the protein. Indeed, exomes, while constituting only 1% of the genome, are the location of 85% of pathogenetic mutations. The number of mutations known to be associated with a genetic disease exceeds 110,000 variants, in most of 3700 different genes. It was estimated that only half of the Mendelian diseases have a known genetic basis. For these reasons, the scientific research focused mostly on exomes for the identification of new genes [21]. The approach, commonly used in the past, was to identify the transmitted loci associated with the phenotype by means of the segregation or linkage analysis, identifying shared genes of affected individuals in large size families. The advent of WES technology has

triggered a rapid growth of the sector (identification of new disease), having the advantage of requiring a limited number of samples. The first successful application of this method has led to the identification of the gene DHODH as causative of Miller syndrome. Since 2010, a large amount of studies have identified new disease-causing mutations, through the exome sequencing method using a reduced number of affected individuals, in various types of diseases including cases of neuropathy, poikiloderma associated with neutropenia, familial exudative vitreoretinopathy, immune disorders, and tumor predisposition. The use of WES allowed to identify some causative mutations also in diseases with high phenotypic heterogeneity where the traditional linkage analysis approach is more difficult. A new area of the development of NGS technology is linked to the identification of new biomarkers or pharmacogenomics with the development of personalized therapies. The impact of next-generation sequencing technology on genomics has in turn led to a new revolution in the genetic field changing the nature of genetic trials. The production of a large number of low cost NGS platforms is useful for many applications. These include: new variants discovered by targeted resequencing regions of interest or whole genomes; *de novo* assemblies of bacterial genomes and lower eukaryotes; cataloging of cell, tissues and organism transcriptomes (RNA-Seq); profiling of genome-wide epigenetic marks and chromatin structure using seq-based methods (ChIP-seq and Methyl-seq); classification of species and/or discovery of the gene using metagenomic studies. NGS technologies can also accelerate exploration of the natural world [22, 23]. Despite a dramatic increase in the number of complete genome sequences available in public databases, most of the biological diversity, in our world, remains to be explored. Nowadays, *de novo* assembly of NGS data requires the development of new software tools that can overcome the technical limitations of these technologies. In fact, the main limitation is a rapid deterioration in the quality of assembly as the length of reading decreases.

#### 4.3.8. *De novo* assembly

*De novo* genome assembly is often compared to solving a great puzzle without knowing the picture that we are trying to rebuild [7, 21, 24]. Mathematically, the issue *de novo* assembly is difficult regardless of the method of sequencing. During the *de novo* assembly process, a high number of repeated segments in the genome may cause several errors. The assembler tool should guess the right genome, starting from a great number of alternative options (the number of attempts increases exponentially with a high pattern of repetitions in the genome) [25]. As the technology has evolved, new methods for assembling genomes have continuously changed [26]. Genome sequencers have never been able to read more than a relatively short stretch of DNA at once, with read lengths gradually increasing over time.

**Assembly quality:** High coverage is necessary to sequence polymorphic alleles within diploid or polyploid genomes. However, using shorter reads, the coverage should be increased in order to balance the low connectivity of the system and to obtain an optimal assembly. However, some times, a poor assembly process cannot be improved by a higher coverage. In fact, in the case of a high number of long repeated sequences in the genome, a high coverage will never fully compensate for the increasing errors and gaps which are produced during the editing phase. Using paired end technology gaps can be crossed and eliminated [5].

Assembly methods aim to create the most comprehensive reconstruction as possible without introducing errors. The central challenge of assembling the genome is to solve the problem of repetitive sequences. If the DNA reads are random, then the expected number of cases of each sequence will decrease exponentially and the number of repeats in the genome is reduced. However, several genomes may share highly repeated structures which do not allow an easy assembly process [27].

**Scaffolding:** The phase of scaffolding focuses on repeats to fix linking contig initials in data-driven scaffolds. A scaffold is a collection of contig linked by pair partners, where gaps between contigs between contig might be repeats, in which the gap can in theory be filled or outright gaps that original sequencing project does not capture [6]. If the distances are long enough, the assembler is able to connect contig in nearly all repetitions. Assembler tools may vary in the method how they call the contigs. The great majority of them are based on a combination of two factors: contig length and a number of reads. A contig containing too many reads is called as a repetition. High-copy-number patterns are easy to be identified. On the contrary, the identification of two copies is more difficult) [25]. If the contigs are overlapping in a scaffold, the assembler can merge at this point. Otherwise, the assembler will record a gap inside the scaffold.

*Example of de novo assembly using the galaxy platform ([www.galaxy.org](http://www.galaxy.org))*

*(a) Read quality control*

Steps involved and suggested tools:

*NGS: QC and manipulation → FastQC: comprehensive QC*

*Command line: fastqc*

Some of the important outputs of FastQC for our purposes are:

- Read length: Will be important in setting maximum k-mer size value for assembly;
- Quality encoding type: Important for quality trimming software;
- % GC: High GC organisms don't tend to assemble well and may have an uneven read coverage distribution;
- Total number of reads: Gives you an idea of coverage;
- Dips in quality near the beginning, middle, or end of the reads: Determines possible trimming/cleanup methods and parameters and may indicate technical problems with the sequencing process/machine run;
- Presence of highly recurring k-mers: May point to contamination of reads with barcodes, adapter sequences, etc.;
- Presence of large numbers of Ns in reads: May point to poor quality sequencing run. You need to trim these reads to remove Ns.

*(b) Quality trimming/cleanup of read files*

Steps involved and suggested tools:

*NGS: QC and manipulation > trimmomatic*

Command lines:

- Adapter trimming
- Sliding window trimming
- Trailing bases quality trimming
- Minimum read length

*(c) Assembly*

Steps involved and suggested tools:

*NGS-Assembly → Velvet Optimizer*

*Possible alternative software:*

*Spades*

*SOAP-denovo*

*MIRA*

*ALLPATHS*

*Possible tools for improving your assemblies:*

*QUAST – <http://bioinf.spbau.ru/quast>*

*Mauve*

*InGAP-SV – <https://sites.google.com/site/nextgengenomics/ingap>*

*Semi-automated gap fillers:*

*Gap filler – <http://www.baseclear.com/landingpages/basetools-a-wide-range-of-bioinformatics-solutions/gapfiller/>*

*IMAGE – <http://sourceforge.net/apps/mediawiki/image2/>*

*Genome visualizers and editors:*

*Artemis – <http://www.sanger.ac.uk/resources/software/artemis/>*

*IGV – <http://www.broadinstitute.org/igv/>*

*Geneious – <http://www.geneious.com/>*

*CLC BioWorkbench – <http://www.clcbio.com/products/clc-genomics-workbench/>*

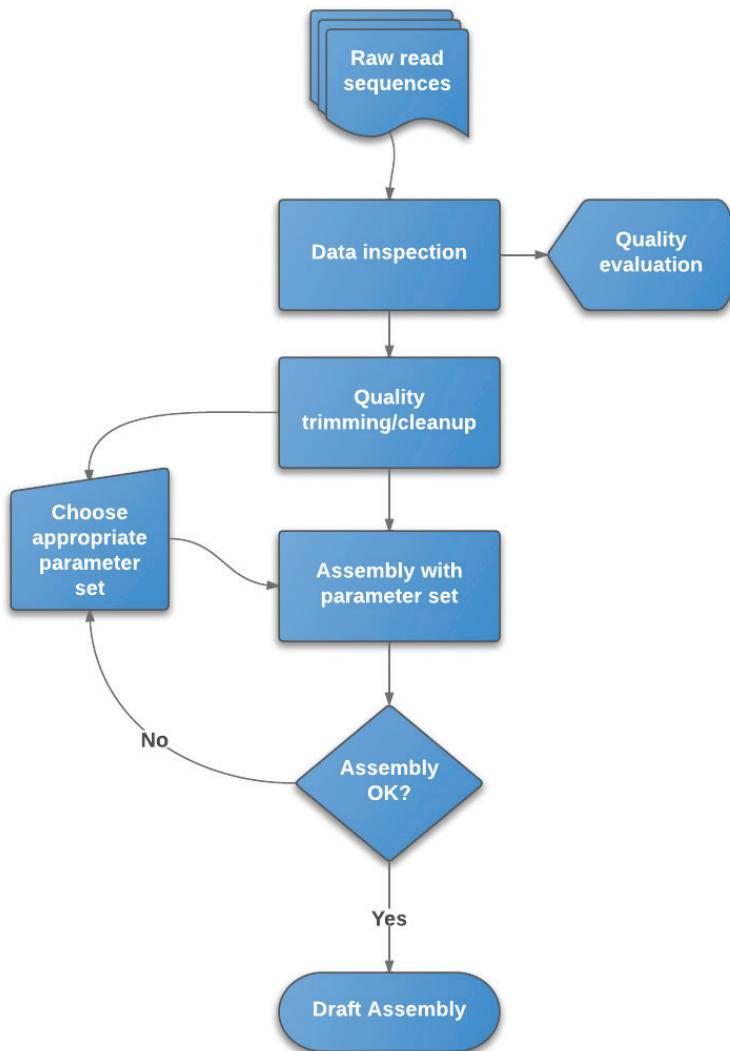
*Automated and semi-automated annotation tools:*

Prokka – <https://github.com/tseemann/prokka>

RAST – <http://www.ncbi.nlm.nih.gov/FIG/wiki/view.cgi/FIG/RapidAnnotationServer>

JCVI – <http://www.jcvi.org/cms/research/projects/annotation-service/>.

**Figure 2** shows an example of a genome assembly workflow.



**Figure 2.** Example of a genome assembly workflow.

#### 4.3.9. RNA-seq method

The transcriptome is the totality of the transcripts present in a cell, each with their abundance and varies depending on the age stage and physiological conditions of the animals. The study of the transcriptome allow to interpret the functional genome elements identifying the different types of transcripts (mRNA, coding RNA, and small RNAs), to determine their level of expression and their structure (terminations to the 5' and 3', exon-intron structure, alternative splicing). The introduction of NGS methods has revolutionized the characterization and quantification of transcriptomes overcoming limitations of previous methods (for example, microarrays or RT-PCR), as the need for a previous knowledge of the genomic sequence, the background noise (due to cross-hybridization), and the limited dynamic range. Starting from the mRNA, a cDNA library is constructed and specific adapters to both ends of each sequence are used. Each molecule is then sequenced after the PCR amplification. The sequencing produces a series of short reads, which can be aligned to the genome or the reference transcriptome. These are used to construct an expression profile for each gene, with a resolution that can reach up to a single base. If the reads are short and very numerous (up to several million per run) you can identify transcripts with a low level of expression. The RNA-seq is highly reproducible and requires less RNA for the synthesis library, since it lacks the step of cloning. Nc RNAs (noncoding RNA) are small molecules which are not translated into proteins. This class of RNAs includes several RNAs as the transfer RNA (tRNA), ribosomal RNA (rRNA), small nuclear and small nucleolar RNA, micro RNA and small interfering RNA (miRNA and siRNA). Micro RNAs (21-nucleotide-long RNAs) are very important posttranscriptional regulators of genes in animals. **Table 3** reports the most important Internet reference databases.

NCBI-RefSeq	<a href="http://www.ncbi.nlm.nih.gov/refseq">http://www.ncbi.nlm.nih.gov/refseq</a>
UCSC	<a href="http://genome.ucsc.edu">http://genome.ucsc.edu</a>
Ensembl	<a href="http://www.ensembl.org/index.html">http://www.ensembl.org/index.html</a>
dbSNP	<a href="http://www.ncbi.nlm.nih.gov/SNP/index.html">http://www.ncbi.nlm.nih.gov/SNP/index.html</a>
dbVar	<a href="http://www.ncbi.nlm.nih.gov/dbvar">http://www.ncbi.nlm.nih.gov/dbvar</a>
	<a href="http://ncbi.nlm.nih.gov/dbvar/content/overview">http://ncbi.nlm.nih.gov/dbvar/content/overview</a>
dbGaP	<a href="http://www.ncbi.nlm.nih.gov/gap">www.ncbi.nlm.nih.gov/gap</a>
DNA Data of Japan	<a href="http://www.ddbj.nig.ac.jp">http://www.ddbj.nig.ac.jp</a>
ExPASy	<a href="http://ca.expasy.org">http://ca.expasy.org</a>
GenAtlas	<a href="http://citi2.fr/GENATLAS/">http://citi2.fr/GENATLAS/</a>
GenBamk	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>
SNP consortium	<a href="http://snp.ncbi.nlm.nih.gov">http://snp.ncbi.nlm.nih.gov</a>
Stanford microarray	<a href="http://genome.stanford.edu/microarray">http://genome.stanford.edu/microarray</a>
Swiss-Protein	<a href="http://www.expasy.org/sprot">http://www.expasy.org/sprot</a>

**Table 3.** List of reference databases (Internet sites).

## 5. Other methods for screening the genetic diseases

### 5.1. Mendelian disorders

Karyotyping and *in situ* fluorescent hybridization technique (FISH) were traditionally used to detect the large genetic disorders. In the last decade, microarrays have been used to improve the resolution. Nowadays, NGS technologies can be used to sequence all the exomes or targeted exome regions.

### 5.2. Microarrays or SNP chip

SNP array is a specific type of chip array which is used to identify the single point mutations. After the DNA extraction, DNA is fragmented and processed by means of biochemical methods and labeled with a fluorescent dye. The DNA spots (20–100 bp) are attached to a solid surface (glass, plastic, or silicone) and the hybridization is performed using the samples of DNA to be tested. For each SNP to be genotyped, several DNA probes are designed with similar sequences but variable at one position corresponding to a polymorphism. The measure of the fluorescence, at each spot, allows us to detect the presence/absence of a specific allele. Two individual are identical at 99.9% of the genome and the wide linkage disequilibrium (LD) observed throughout the genome reduces the number of needed SNPs. If many SNPs are linked together in one region, only one marker is needed to be accurately measured. The ability to detect SNPs using array-based approaches is often limited by the density of the array.

### 5.3. Epigenetics

Methylation effects and histone modifications explain most of the epigenetic and posttranscriptional modifications. Chip-seq (chromatin-immuno-precipitation and direct sequencing) is used at the whole-genome level [28].

### 5.4. Analysis of a nucleotide sequence

The operations to be carried out in the molecular laboratory are as follows:

- extracting DNA from a tissue, usually the blood;
- DNA amplification. It consists in amplifying selectively the DNA;
- DNA sequencing. Once you obtained the various DNA fragments amplified by PCR, you run the “sequencing reaction”;
- Analysis of the nucleotide sequences and search for gene mutations.

Specific software is used to perform an alignment between the nucleotide sequences of an individual showing a monogenic disease (from which you want to search for the mutation) and those of a control healthy individual (i.e., which does not have one specific disease). From this comparison, if the animal has a monogenic disease, may emerge one or more differences in the nucleotide sequence. These mutations (silent mutations, missense, not sense, frameshift, and splicing) shall be individually analyzed and evaluate the effects. The prediction of the effects of a missense mutation in the protein structure level is a

complex process, in which we must take into account all the characteristics of the amino acids protein chain and their mutual interactions. Currently, some programs are available (such as SwissPdbViewer, for free on the website <http://www.expasy.org/spdbv/text/download.htm>), able to process all of these data and provide a prediction of how the mutation will affect the three-dimensional structure of the protein. The automated sequencer provides the data files that can be displayed in the form of electropherograms with specific programs such as MT-Navigator (Applied Biosystems). The succession of peaks that make up the electropherogram is the result of the detection the fluorescence emitted by the fluorophores linked to the four ddNTPs used in the reactions of sequence. For each fluorochrome, and then to each nucleotide, it is conventionally associated a different color (A = green, T = red, C = blue, and G = black), thus facilitating the reading of the sequence. The reference sequences for the comparison may be obtained from the National Database Center for Biotechnology Information (NCBI, website: <http://www.ncbi.nlm.nih.gov>). The individual mutations are recognizable on the basis of several peak patterns. The substitutions of a single nucleotide or more nucleotides can be recognized because of its presence in a specific point of different color peaks corresponding to the replaced bases. One individual homozygous shows only one peak while a heterozygous shows two peaks corresponding to two alleles. A mutation can be localized in correspondence of the first base of the junction of splicing (GT becomes TT). In fact in most eukaryotic genes, introns begin with the "GT" bases and end with "AG."

## 6. Clinical applications in domestic animals

Animal geneticist, whose role consisted only once in calculating the probability of transmission hereditary of a given genetic disease in a given individual, is now an active figure throughout the clinical process, from diagnosis to treatment. The collection of family history is a central point in clinical genetics. It, in fact, can help to arrive at a correct diagnosis. In addition, through early diagnosis, it is possible to prevent or delay the onset of the disease through a suitable therapy. This possibility is often about relatives of an individual and can take to extend the analysis beyond the individual subject. Even for a correct formulation of the prognosis, it is important to be able to refer to the specific form assumed by a given disease in a certain family. For example, in some cases, analysis of family history has allowed to recognize some individuals with less aggressive forms of a disease, a less severe clinical picture, and then improved life expectancy.

## 7. Genetic counseling

Genetic counseling is a communication process that aims to provide to individual pet owners or farmers the risk of genetic diseases in order to help these people to analyze the consequences and to make responsible decisions. The first task of the genetic counselor is to identify the individual's needs and provide information with an appropriate manner to the degree of its culture and its beliefs. In addition to information about the disease, it must

present the possible available genetic tests, highlighting their reliability. Finally, it should discuss the implications of the various choices.

## Author details

Edo D'Agaro

Address all correspondence to: [edo.dagaro@uniud.it](mailto:edo.dagaro@uniud.it)

Department of Agricultural, Food, Environment and Animal Sciences, University of Udine, Udine, Italy

## References

- [1] Mardis, E R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* 2008; **24**: 133–141.
- [2] Schatz M C, Delcher A L, Salzberg, S L. Assembly of large genomes using second – generation sequencing. *Genome Res.* 2010; **20**: 1165–1173.
- [3] Dong Y. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). *Nat. Biotechnol.* 2013; **31**: 135–141.
- [4] Narzisi G, Mishra B, Schatz MC. Algorithms for computational biology. *Lecture Notes in Computer Science.* 2014; **21**: 183–195.
- [5] Gnerre S. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. USA.* 2011; **108**: 1513–1518.
- [6] Berlin K. Assembling large genomes with single-molecule sequencing and locality – sensitive hashing. *Nat. Biotechnol.* 2015; **33**: 623–633.
- [7] Chaisson, M J, et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature.* 2015; **517**: 608–611.
- [8] Pendleton M. Assembly and diploid architecture of an individual human genome via single-molecule technologies. *Nat. Methods.* 2015; **12**: 780–786.
- [9] Sharon, D., Tilgner, H., Grubert, F., Snyder, M. A single-molecule long-read survey of the human transcriptome. *Nat Biotechnol.* 2013; **31**: 1009–1014.
- [10] Gordon D. Long-read sequence assembly of the gorilla genome. *Science.* 2016; **352**: aae0344.
- [11] Koren S. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol.* 2013; **14**: R101.
- [12] Sebat J. Large-scale copy number polymorphism in the human genome. *Science.* 2004; **305**: 525–528.

- [13] Browning S R, Browning B L. Haplotype phasing: existing methods and new developments. *Nat. Rev. Genet.* 2011; **12**: 703–714.
- [14] Zheng G X. Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat. Biotechnol.* 2016; **34**: 303–311 (2016).
- [15] Phillippy A M, Schatz M C, Pop M. Genome assembly forensics: finding the elusive mis-assembly. *Genome Biol.* 2008; **9**: R55.
- [16] Baker M. Structural variation: the genome's hidden architecture. *Nat. Methods.* 2012; **9**: 133–137.
- [17] Hosomichi K, Shiina T, Tajima A, Inoue I. The impact of next-generation sequencing technologies on HLA research. *J. Hum. Genet.* 2015; **60**: 665–673.
- [18] Kuleshov V. Whole-genome haplotyping using long reads and statistical methods. *Nat. Biotechnol.* 2014; **32**: 261–266.
- [19] Chen X. The architecture of a scrambled genome reveals massive levels of genomic rearrangement during development. *Cell.* 2014; **158**: 1187–1198.
- [20] Ross M G. Characterizing and measuring bias in sequence data. *Genome Biol.* 2014; **14**: R51.
- [21] Cao H. Rapid detection of structural variation in a human genome using nanochannel-based genome mapping technology. *GigaScience.* 2014; **3**: 34.
- [22] Oulas A. Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinform. Biol. Insights.* 2015; **9**: 75–88.
- [23] Voskoboinik A. The genome sequence of the colonial chordate, *Botryllus schlosseri*. *eLife.* 2013; **2**: e00569.
- [24] Li R. The sequence and de novo assembly of the giant panda genome. *Nature.* 2010; **63**: 311–317.
- [25] Church D M. Extending reference assembly models. *Genome Biol.* 2016; **16**: 13.
- [26] Putnam NH. Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res.* 2016; **26**: 342–350.
- [27] Paszkiewicz K, Studholme DJ. *De novo* assembly of short sequence reads. *Brief Bioinform.* 2010; **11**(5):457–72.
- [28] Fang G. Genome-wide mapping of methylated adenine residues in pathogenic *Escherichia coli* using single-molecule real-time sequencing. *Nat. Biotechnol.* 2012; **30**: 1232–1239.

