Review Article

The Application of Genomic Technologies to Investigate the Inheritance of Economically Important Traits in Goats

Marcel Amills

Department of Animal Genetics, Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Campus Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Correspondence should be addressed to Marcel Amills; marcel.amills@uab.es

Received 13 February 2014; Revised 15 July 2014; Accepted 19 July 2014; Published 17 September 2014

Academic Editor: Guillermo Giovambattista

Copyright © 2014 Marcel Amills. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Goat genomics has evolved at a low pace because of a lack of molecular tools and sufficient investment. Whilst thousands and hundreds of quantitative trait loci (QTL) have been identified in cattle and sheep, respectively, about nine genome scans have been performed in goats dealing with traits as conformation, growth, fiber quality, resistance to nematodes, and milk yield and composition. In contrast, a great effort has been devoted to the characterization of candidate genes and their association with milk, meat, and reproduction phenotypes. In this regard, causal mutations have been identified in the \( \alpha_S^1 \)-casein gene that has a strong effect on milk composition and the PIS locus that is linked to intersexuality and polledness. In recent times, the development of massive parallel sequencing technologies has allowed to build a reference genome for goats as well as to monitor the expression of mRNAs and microRNAs in a broad array of tissues and experimental conditions. Besides, the recent design of a 52K SNP chip is expected to have a broad impact in the analysis of the genetic architecture of traits of economic interest as well as in the study of the population structure of goats at a worldwide scale.

1. Introduction

The main purpose of this review is to provide a general perspective of the advances made in the field of goat genomics in the last three decades. Goats are a species with a lower economic value than other domesticates, as cattle and pigs. They are mainly raised in Asian (\( \sim \)500 million heads) and African (\( \sim \)290 million heads) countries, whilst their relevance in Europe (\( \sim \)21 million heads) and North America (\( \sim \)3 million heads) is relatively modest [1]. These circumstances may explain why the genetic study of goats has experienced, in general, a substantial lag behind those performed in bovines and even sheep, a closely related species. Whilst the genomics analysis of quantitative traits has undergone substantial advances in the two species mentioned before, leading to quite a few successful cases to the identification of causal mutations, a small number of studies have identified quantitative trait loci (QTL) in goats. Fortunately, there are compelling signs that this situation is about to change, mainly because of the development of high throughput genotyping and sequencing tools that are allowing to generate huge amounts of data with a moderate investment of time and money. In the following pages, an outline of the major findings in the genetic analysis of quantitative and Mendelian traits of economic interest will be provided. Next, the impact of massive genotyping and sequencing platforms on the characterization of the caprine genome and transcriptome will be discussed. The review will conclude with some comments about future trends and developments in the field of goat genomics.

2. Analysing the Genetic Architecture of Production and Disease-Related Traits in Goats

The genetic analysis of production and disease-related traits in goats has been rarely done at a genome-wide scale. In this regard, the lack of well-established microsatellite panels covering the whole genome hindered, to a significant extent, the implementation of genome scans aimed at detecting QTL. Whilst 8,305 and 789 QTL have been reported...
**Table 1: Identification of quantitative trait loci for economically important traits in goats.**

<table>
<thead>
<tr>
<th>Type of trait</th>
<th>Phenotype</th>
<th>Chromosome</th>
<th>Population</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conformation</strong></td>
<td>Head length</td>
<td>1, 4</td>
<td>Angora</td>
<td>Marrube et al. (2007) [5]</td>
</tr>
<tr>
<td></td>
<td>Body length</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chest depth</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chest circumference</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td>Body weight</td>
<td>4, 8, 17, 27</td>
<td>Angora</td>
<td>Visser et al. (2013) [6]</td>
</tr>
<tr>
<td></td>
<td>Weaning weight</td>
<td>16, 19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bodyweight, average daily gain</td>
<td>1, 2, 5</td>
<td>Rayini</td>
<td>Mohammad Abadi et al. (2009) [9]</td>
</tr>
<tr>
<td><strong>Mohair fiber quality and weight</strong></td>
<td>Coeff. var. fiber diameter</td>
<td>1, 13</td>
<td>Angora</td>
<td>Cano et al. (2007) [3]</td>
</tr>
<tr>
<td></td>
<td>Kemp fiber</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discontinuous medullated fibers</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staple length</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fleece weight</td>
<td>2, 5, 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fiber diameter</td>
<td>4, 24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comfort factor Spinning firmness</td>
<td>8, 13, 18, 20</td>
<td>Angora</td>
<td>Visser et al. (2011) [4]</td>
</tr>
<tr>
<td><strong>Cashmere fiber</strong></td>
<td>Cashmere yield</td>
<td>2, 5, 13</td>
<td>Rayini</td>
<td>Mohammad Abadi et al. (2009) [9]</td>
</tr>
<tr>
<td><strong>Resistance to gastrointestinal nematodes</strong></td>
<td>Fecal egg count</td>
<td>22, 26</td>
<td>Creole</td>
<td>De La Chevrotière et al. (2012) [2]</td>
</tr>
<tr>
<td></td>
<td>Eosinophil count</td>
<td>7, 8, 14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Packed cell volume</td>
<td>5, 9, 21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IgE anti-L3</td>
<td>3, 16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IgE anti-ESP</td>
<td>1, 26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Worm egg count</td>
<td>23</td>
<td>Angora, Cashmere</td>
<td>Bolormaa et al. (2010) [7]</td>
</tr>
<tr>
<td><strong>Milk traits</strong></td>
<td>Milk yield</td>
<td>6, 14</td>
<td>—</td>
<td>Roldán et al. (2008) [8]</td>
</tr>
<tr>
<td></td>
<td>Fat percentage</td>
<td>14, 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protein percentage</td>
<td>3, 20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Partial genome scan for chromosome 23 (major histocompatibility complex region).
2. Shape parameter of the lactation curve for each one of the phenotypes.
3. Partial genome scan for 4 chromosomes (3, 6, 14, and 20).

In the Animal QTL Database (http://www.animalgenome.org/QTLdb/release.php) for cattle and sheep, respectively, goats are not even included. A bibliographic search at NCBI Pubmed Database (http://www.ncbi.nlm.nih.gov/) revealed few publications where caprine QTL are described (a representative list can be found at Table 1). For instance, De la Chevrotière et al. [2] identified 13 QTL for resistance to gastrointestinal nematode infections in Creole goats by analysing 101 microsatellites. Quantitative trait loci affecting mohair [3, 4] and conformation traits [5] as well as preweaning growth [6] have also been reported in Angora goats. On the other hand, partial genome scans have been carried out by Bolormaa et al. [7], who described QTL for worm egg and blood eosinophil counts in Australian Angora and Cashmere goats through the analysis of three microsatellites mapping to the caprine major histocompatibility complex on chromosome 23, and by Roldán et al. [8] who genotyped 37 microsatellites mapping to four chromosomes and found several QTL influencing the variation of milk yield and quality traits across lactation. Mohammad Abadi et al. [9] have also identified QTL for growth and Cashmere fiber.

In the absence of positional information, the investigation of the genetic factors that determine phenotypic variation of economically important traits has been based on physiological candidate gene approaches. This strategy has important limitations. The involvement [10] of one gene in a metabolic pathway does not necessarily imply that it contains variation affecting the trait under study. A common pitfall of association studies is to report significance as raw P values, instead of correcting for multiple testing. Other extended flaws are to report associations without proposing any biological mechanism to support them or to infer that a trait is associated with a given genotype on the basis of divergent allelic frequencies among populations with extreme phenotypes. Flawed experimental practices can lead to the publication of spurious associations that do not have any biological basis. The recent opportunity to carry out genome-wide association studies in goats is expected to alleviate this problem by making possible to generate positional information at an affordable cost and relatively high resolution.

2.1. Candidate Genes for Dairy Traits. Candidate genes related with milk traits have been widely studied in goats [11–13]. A detailed description of the polymorphism of casein, β-lactoglobulin, and α-lactalbumin genes can be found at Moioli et al. [11] and Amills et al. [13] and will not be reviewed.
Advances in Biology here. Obviously, genes encoding caseins are strong candidates to explain the variation of traits such as milk protein content, rheological parameters, and cheese yield. The casein cluster maps to a 250–300 kb region on chromosome 6 and consists of four loci, namely, αS1-casein (CSN1S1), β-casein (CSN2), αS2-casein (CSN1S2), and κ-casein (CSN3). Three decades ago, Boulanger et al. [14] identified, by starch gel protein electrophoresis, CSN1S1 variants differing in intensity, an observation that suggested that the polymorphism of this gene might have differential effects on casein synthesis. Rocket immunoelectrophoresis studies provided evidence of the existence of CSN1S1 alleles with quantitative effects on CSN1S1 content [15]. So far, the catalog of CSN1S1 alleles has expanded to a total of 17 variants that can be classified, according to the CSN1S1 content they determine, as strong (A, B1, B2, B3, B4, C, H, L, and M), medium (E and I), low (F, D, and G), and null (01, 02, and N). Importantly, causal mutations explaining these quantitative effects have been identified, providing a valuable model to understand how genetic variation modulates milk composition. In this way, the most distinguishing feature of the E-allele is the presence of a retrotransposon insertion at exon 19 that might destabilize the transcript and shorten its half-life [16]. Similarly, the G-allele of the bovine CSN1S1 gene, which is also associated with a reduced milk CSN1S1 percentage, contains a retrotransposon insertion at exon 19 [17]. On the other hand, low CSN1S1 content alleles contain mutations that perturb the normal splicing of the mRNA, yielding transcripts that encode shorter CSN1S1 proteins because of exon skipping [18–20]. Lastly, a large genomic deletion encompassing intron 12 to exon 19 of the caprine CSN1S1 gene explains the absence of this protein in individuals harbouring the 01 allele [21], while the N-allele is characterized by a premature stop codon at exon 12 [20].

The polymorphism of the goat CSN1S1 gene has broad effects on a wide array of dairy phenotypes. Pioneering studies performed by Mahé et al. [22] and Manfredi et al. [23] revealed that the CSN1S1 genotype influenced milk protein percentage and, more unexpectedly, also fat content. These findings were confirmed by Barbieri et al. [24], who showed a consistent association of the A-allele with higher milk and protein fat contents. Similarly, Zullo et al. [25] reported effects of the CSN1S1 genotype on milk composition in Cilentana goats. Chilliard et al. [26] and Balia et al. [27] have also proposed that milk fat composition is modulated by the CSN1S1 genotype. A global effect of CSN1S1 variation on protein content might be explained by the fact that CSN1S1 plays a pivotal role in casein transportation from the endoplasmic reticulum to the Golgi complex in mammary epithelial cells [28]. Besides, CSN1S1 genotype seems to influence the structure and composition of milk fat globules [29] as well as the expression of lipogenic genes [30]. These data back up the notion that lipid and protein synthesis secretory pathways are tightly interconnected in the mammary gland [31].

Rheological properties of milk are also influenced by the CSN1S1 genotype; for example, milk from AA goats is associated with a firmer curd, a slower coagulation time, and increased cheese yield than that of FF goats in French [32, 33] and Italian breeds [25], whilst in Spanish goats less conclusive results have been obtained [34]. From an organoleptic point of view, the AA cheese has a less pronounced goat flavor intensity than that of the FF one, maybe because of differences in fatty acid composition and lipolysis rate [26, 33].

The variability of the other three casein genes and its association with milk traits have also been explored. Null alleles at the CSN1S2 [35] and CSN2 [36] genes have been described, but to the best of our knowledge association analyses with milk yield and protein and fat contents have not been performed yet. With regard to the CSN3 locus, a high number of missense mutations have been detected by Yahyaoui et al. [37], Jann et al. [38], and Prinzenberg et al. [39], and a nomenclature system for allelic variation at this locus has been proposed. According to Caravaca et al. [40] and Chiatti et al. [41], CSN3 polymorphism is associated with protein and casein contents, and in a recent work Caravaca et al. [34] have also suggested effects on milk rennet coagulation time.

A powerful approach to investigating the role of casein genes in determining milk composition consists in analysing haplotypes rather than specific genes or alleles. Hayes et al. [42] typed 39 SNPs mapping to the caprine casein loci in Norwegian goats and found associations between CSN1S1 haplotypes and protein percentage and fat yield and between CSN3 haplotypic variation and fat and protein percentages. Pazzola et al. [43] have also genotyped the four casein genes in Sarda goats and have described associations with diverse rheological parameters. It is important, however, to emphasize that casein concentrations are not exclusively influenced by polymorphisms within the casein cluster. In cattle, Schopen et al. [44] reported trans-QTL influencing the contents of CSN1S1 (bovine chromosome 9), CSN1S2 (chromosomes 1, 10, and 17), CSN2 (chromosome 3), and total caseins (chromosome 11). This means that genome-wide approaches will be needed to identify the genetic factors that regulate milk casein and protein contents.

There is a growing list of candidate genes whose polymorphism has been associated with milk fat content and composition traits (Table 2). These genes belong to different functional categories related with lipogenesis (ACACA, DGAT1, DGAT2, MEL, and SCD); lipolysis (LPL and LIPE); milk fat globule membrane proteins (BTN1A1 and MFGE); hormone signaling (GH and PRLR); and transcription factors regulating gene expression (PITX2, POUF1, and STAT5), amongst others [45–59]. Although the main phenotype analysed in these reports is milk fat content, in some of them the genetic analysis of milk fatty acid composition has been undertaken (e.g., [53, 55, 58, 59]). However, and in strong contrast with the CSN1S1 model (see above), there is a complete lack of functional studies supporting the associations revealed in each one of these scientific reports. This important limitation hinders the application of this knowledge to improve milk fat composition through marker assisted selection schemes. In the future, high throughput genotyping tools are expected to provide a more comprehensive view about the genetic architecture of milk lipid traits.

2.2. Candidate Genes for Growth and Meat Quality Traits. The characterization of candidate genes influencing growth and meat-related traits has not reached yet the level of knowledge
achieved for milk traits (see previous section), where mutations with causal effects have been identified. Probably, the loci that have been analysed most intensively are those related with the growth hormone (GH) axis. Growth hormone is secreted by the anterior pituitary and its main effects are to stimulate bone and skeletal muscle growth, through the action of IGF1, as well as to increase milk yield and diminish adiposity [60]. A missense SNP at the bovine GH gene has been associated with GH concentrations [61]. Besides, three haplotypes differing by amino acid substitutions at positions 127 and 172 have been associated with carcass weight and beef marbling score in Japanese cattle [62]. In goats, association studies have revealed the existence of relationships between GH genotype and a wide array of growth parameters such as body length and height [63, 64], and birth chest and weaning weight and height [65]. Similarly, the variability of the GH receptor has been associated with body length and height [65], while the growth hormone secretagogue receptor genotype displays significant associations with body length and body length index [66].

Myostatin (MSTN) belongs to the transforming growth factor-β (TGF-β) superfamily and it has been shown to repress muscular growth [67]. In cattle, mutations inactivating MSTN expression lead to a muscular hypertrophy phenotype known as double muscling [68]. In goats, variation at the MSTN gene has been associated with body weight, length, and height [69]. In the case of GH and MSTN, genetic variability may affect growth rate because both molecules are known to play a key role in this physiological process (although causal mutations have not been identified yet). For other loci, this relationship is less obvious; for example, the polymorphism of the diacylglycerol acyltransferase 2 (DGAT2) gene that regulates triacylglycerol synthesis has been associated with withers height in Chinese breeds without providing any mechanistic explanation [70].

### 2.3. Candidate Genes for Reproduction Traits

In Chinese goat breeds, the association between litter size and a wide array of polymorphisms mapping to the BMP4 [71], CART [72], GDF9 [73, 74], GNRH1 [74], INHA and INHBA [75], KISS1 [76–78], KITLG [79], POUF1 [80], and TSHB [81] genes, amongst others, has been screened. However, the most significant finding in the field of caprine reproduction genetics has been the elucidation of the molecular basis of the goat polled intersex syndrome (PIS). This syndrome was reported for the first time by Bourdelle [82], and it is associated with the absence of horns, in males and females, and sex reversal (i.e., masculinization) of females [83]. While the inheritance of polledness is dominant, intersexuality segregates as a recessive trait [84]. The PIS locus was mapped to a 100 kb region on the goat 1q43 chromosome [85]. In humans, this genomic location contains the blepharophimosis ptosis epicanthus inversus syndrome locus, related with eyelid malformation and the loss of ovarian function. Bacterial artificial chromosome sequencing revealed that the causal mutation of caprine PIS is a 11.7 kb deletion affecting the expression of PISRT1, a long noncoding RNA, and FOXL2, a forkhead transcription factor involved in ovarian development and the maintenance of granulosa cell function [86]. The deleted region may contain a long range regulatory element affecting the expression of these two loci that lie 20 kb (PISRT1) and 200–300 kb (FOXL2) apart from it. Expression analyses have shown that, in XX PIS−/− sex-reversed gonads, FOXL2 mRNA levels are greatly diminished as early as 36 days after conception, whilst PISRT1 RNA declines a little bit later [86]. At 56 days after conception and afterwards, the transcript

**Table 2:** List of candidate genes associated milk fat content and composition traits.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Associated phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACACA</td>
<td>Fat yield, lactose content, and somatic cell count</td>
<td>Badaoui et al. (2007) [45]</td>
</tr>
<tr>
<td>BTNIA1</td>
<td>Milk fat yield, total solid, solid-nonfat and first milk yield</td>
<td>Qu et al. (2011) [48]</td>
</tr>
<tr>
<td>DGAT1</td>
<td>Milk yield and fat content</td>
<td>An et al. (2012) [49]</td>
</tr>
<tr>
<td>DGAT2</td>
<td>Milk yield and fat content</td>
<td>An et al. (2011) [50]</td>
</tr>
<tr>
<td>GH</td>
<td>Milk fat and milk fat and protein percentages</td>
<td>Malveiro et al. (2001) [51]</td>
</tr>
<tr>
<td>LIPE</td>
<td>Milk FA composition and fat content</td>
<td>Dettori et al. (2013) [52]</td>
</tr>
<tr>
<td>LPL</td>
<td>Milk fat content</td>
<td>Zidi et al. (2010) [53]</td>
</tr>
<tr>
<td>MEI</td>
<td>Milk FA composition</td>
<td>Zidi et al. (2010) [54]</td>
</tr>
<tr>
<td>MFGE8</td>
<td>Milk fat yield and total solid</td>
<td>Qu et al. (2011) [55]</td>
</tr>
<tr>
<td>PIFX2</td>
<td>Milk fat, lactose and solid content, and second lactation milk yield</td>
<td>Zhao et al. (2013) [56]</td>
</tr>
<tr>
<td>POUF1</td>
<td>Milk yield and milk fat and protein percentage</td>
<td>Daga et al. (2013) [57]</td>
</tr>
<tr>
<td>PRLR</td>
<td>Milk FA composition</td>
<td>Zidi et al. (2010) [58]</td>
</tr>
<tr>
<td>SCD1</td>
<td>Milk FA composition</td>
<td>Zidi et al. (2010) [59]</td>
</tr>
<tr>
<td>STAT5</td>
<td>Milk yield and fat content</td>
<td>An et al. (2012) [47]</td>
</tr>
</tbody>
</table>

2.3. Candidate Genes for Reproduction Traits. In Chinese goat breeds, the association between litter size and a wide array of polymorphisms mapping to the BMP4 [71], CART [72], GDF9 [73, 74], GNRH1 [74], INHA and INHBA [75], KISS1 [76–78], KITLG [79], POUF1 [80], and TSHB [81] genes, amongst others, has been screened. However, the most significant finding in the field of caprine reproduction genetics has been the elucidation of the molecular basis of the goat polled intersex syndrome (PIS). This syndrome was reported for the first time by Bourdelle [82], and it is associated with the absence of horns, in males and females, and sex reversal (i.e., masculinization) of females [83]. While the inheritance of polledness is dominant, intersexuality segregates as a recessive trait [84]. The PIS locus was mapped to a 100 kb region on the goat 1q43 chromosome [85]. In humans, this genomic location contains the blepharophimosis ptosis epicanthus inversus syndrome locus, related with eyelid malformation and the loss of ovarian function. Bacterial artificial chromosome sequencing revealed that the causal mutation of caprine PIS is a 11.7 kb deletion affecting the expression of PISRT1, a long noncoding RNA, and FOXL2, a forkhead transcription factor involved in ovarian development and the maintenance of granulosa cell function [86]. The deleted region may contain a long range regulatory element affecting the expression of these two loci that lie 20 kb (PISRT1) and 200–300 kb (FOXL2) apart from it. Expression analyses have shown that, in XX PIS−/− sex-reversed gonads, FOXL2 mRNA levels are greatly diminished as early as 36 days after conception, whilst PISRT1 RNA declines a little bit later [86]. At 56 days after conception and afterwards, the transcript
levels of these two loci in the gonads are undetectable. In contrast, in the horn buds of PIS−/− and PIS+/− 70-day-old fetuses FOXL2 and PISRT1 mRNA levels are strongly increased [86]. These molecular findings agree well with the dominant and recessive inheritance patterns observed for intersexuality (loss of function) and polledness (gain of function). Abolished expression of FOXL2 results in the activation of the testis differentiation program and in the reduced transcription of the CYP19 gene, which converts androgens into estrogens, while PISRT1 RNA has been hypothesized to act as an inhibitor of male differentiation genes as SOX9 [87].

2.4. Candidate Genes for Coat Color. Goats display a wide array of pigmentation patterns, even at the within-breed level (Figure 1), that contrast strongly with the monochromous brown color of the bezoar. The development of multiple coat colors has been probably the result of human artificial selection choosing certain pigmentation phenotypes because of cultural, religious, or breeding practices. Our understanding of the genetic basis of coat color in goats is more limited than that of other species, such as cattle and sheep, where causal mutations with well-established effects have been identified [88]. Pigmentation is a polygenic trait influenced by the combined action of many genes that often interact in an epistatic mode [89].

Although the list of candidate genes for pigmentation patterns can be very large, there are certain loci that seem to have a prominent and consistent role across species. An example of this would be the melanocortin 1 receptor (MC1R) that encodes a G protein-coupled receptor with 7 transmembrane-spanning domains. The activity of this protein determines if either eumelanin (black) or pheomelanin (red/yellow) will be synthesized within the melanocyte [90].

In cattle, dominant black (L99P) and recessive red (premature stop codon) coat colors are produced by alleles at the MCIR gene ([91, 92], reviewed by [93]), and in sheep a M73K mutation is associated with dominant black [94]. In goats, Fontanesi et al. [95] characterized the diversity of the MC1R gene and found one nonsense mutation (p.Q225X), three missense mutations (p.A81V, p.F250V, and p.C267W), and one silent mutation. The Q225X polymorphism is expected to abolish MC1R function but, surprisingly, it did not show a complete association with a pheomelanic pigmentation, maybe because epistatic/modifier genes influence its effects. In Murciano-Granadina goats, the p.267W allele was present in black individuals but not in the brown ones suggesting a causal effect [95].

The involvement of the Agouti (ASIP) locus, which encodes a paracrine hormone that upon binding MCIR reduces its basal activity and stimulates the synthesis of pheomelanin [90], was highlighted by Adalsteinsson et al. [96], who reported the existence of 11 alleles, that is, white or tan (Awt), black mask (Abm), bezoar (Abz), badger face (Ab), grey (Ag), light belly (Alb), Swiss markings (Aam), lateral stripes (Als), mahogany (Amh), red cheek (Aac), and nonagouti (Aa). Fontanesi et al. [97] analysed the variation of ASIP and identified three missense polymorphisms mapping to conserved positions of the cysteine-rich carboxy-terminal domain of the protein (p.A96G, p.C126G, and p.V128G), but they were not completely associated with pigmentation patterns. Besides, Badaoui et al. [98, 99] have characterized the polymorphism of the caprine ASIP and TYRPI genes in Spanish and Italian breeds, whilst Adefenwa et al. [100] reported ASIP variability in Nigerian goats. Interestingly, Fontanesi et al. [97] also found evidence of a copy number variation affecting the ASIP gene that may correspond to the
Table 3: Hereditary diseases in goats where causal mutations have been identified (Online Mendelian Inheritance Database in Animals, http://omia.angis.org.au/home/).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Gene</th>
<th>Causal mutation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-mannosidosis</td>
<td>MANBA</td>
<td>1398delG, leads to frame shift and premature stop codon</td>
<td>Leipprandt et al. (1996) [117]</td>
</tr>
<tr>
<td>Goiter</td>
<td>TG</td>
<td>C to G substitution that produces a stop codon at position 296</td>
<td>Veenboer and de Vijlder (1993) [118]</td>
</tr>
<tr>
<td>Mucopolysaccharidosis IIID</td>
<td>GNS</td>
<td>Nonsense mutation in codon 102</td>
<td>Cavanagh et al. (1995) [119]</td>
</tr>
<tr>
<td>Myotonia</td>
<td>CLCN1</td>
<td>Ala to Pro substitution at 885</td>
<td>Beck et al. (1996) [120]</td>
</tr>
</tbody>
</table>

white/tan dominant A\(^{wt}\) allele. This observation agrees well with data obtained in sheep demonstrating that the white dominant coat is caused by a 190 kb tandem duplication, encompassing the ovine ASIP and AHCY coding regions and the ITCH promoter region [101].

2.5. Genetic Determinants of Disease Susceptibility. Susceptibility to monogenic and complex diseases has a strong impact on the economic output of goat farms. Disentangling the genetic factors that modulate disease progression might be useful to implement selection schemes aimed at eradicating or decreasing the incidence of pathological conditions (for a review see [102]). One of the main determinants of the immune response elicited against pathogens is the major histocompatibility complex (MHC), which encodes class I and II molecules that present antigens to CD8\(^{+}\) and CD4\(^{+}\) T cells, respectively.

In goats, the MHC maps to chromosome 23, encompasses around 2.4 Mb, and contains 160 protein-coding genes [103]. At least two MHC class I loci have been identified by Zidi et al. [104], whilst Dong et al. [103] sequenced the whole MHC region and found four class I genes. Besides, the class II region contains DR [105–107] and DQ genes [108, 109]. As in other domestic species, the DRB gene is extraordinarily polymorphic in goats [106, 107, 110]. One particularity of ruminants, when compared to humans, is that DQ genes are duplicated (reviewed in [111]). The variability of goat MHC class I proteins has been linked to resistance to the caprine arthritis-encephalitis virus infection, a disease that causes joint inflammation and lameness in adults and encephalitis in kids [112]. Other diseases where MHC polymorphisms may play an important role in the development of an appropriate immune response are cowdriosis [113, 114], trichostrongyliasis [115], and Johne’s disease [116]. It should be emphasized, however, that the immunological basis of these associations has not been elucidated yet.

With regard to genetic disorders, a screening of the Online Mendelian Inheritance Database in Animals (http://omia.angis.org.au/home/) reveals the existence of 74 phenotypes which may have genetic causes in goats, including acondroplasia, anophthalmia, cryptorchidism, epidermolysis bullosa, hemimelia, diaphragmatic hernia, hypospadias, Legg-Calvé-Perthes disease, congenital myopathy, lipofuscinosis, peromelia, and many other pathological conditions. Unfortunately, causal mutations have been identified only for a few Mendelian genetic disorders (Table 3, [117–121]).

Transmissible spongiform encephalopathies are a group of fatal neurological diseases including scrapie in sheep and goats and spongiform encephalopathy in cattle. The main pathogenic event is the conformational conversion of the normal prion PrP\(^{c}\) into the PrP\(^{Sc}\) amyloidogenic and misfolded isofrom that tends to form aggregates with neurotoxic effects [122, 123]. In sheep, at least 10 mutations in the PrP\(^{c}\) gene have been associated with resistance/susceptibility to classical scrapie while two others have been linked to atypical scrapie [124, 125]. The goat PrP\(^{c}\) gene is very polymorphic with 28 missense substitutions described so far [126]. Recently, Goldmann et al. [127] have reported that the Met142 residue is associated with increased resistance to preclinical and clinical scrapie, whilst the Ser127 substitution is linked to a lower risk of showing clinical signs of scrapie in goats in which PrP\(^{Sc}\) has accumulated in brain or periphery.

Analysis of gene expression with bovine microarrays has been useful in identifying the genes that are upregulated or downregulated in response to a challenge with Staphylococcus aureus in the mammary gland [128]. Among the genes with an increased expression, it is worthy to mention those related to the establishment of an immune and inflammatory response as well as with the regulation of innate resistance to pathogens and cell metabolism. These loci might constitute the initial line of defense against udder pathogens. Interestingly, a downregulation of lipid metabolism genes was also observed, a feature that is consistent with the inhibition of this biochemical pathway as a consequence of intramammary infection [128].

3. New Tools for Analysing the Goat Genome and Transcriptome

Although the analysis of the goat genome and transcriptome at a large scale is still in its infancy, some important milestones have been already reached. Without a doubt, the most significant of these advances has been the sequencing of a ∼2.66 Gb genome of a 3-year-old female Yunnan black goat by Dong et al. [103], in the framework of the International Goat Genome Consortium (http://www.goatgenome.org/). These authors used a double strategy, based on short-read sequencing and optical mapping that facilitated in a very significant way genome assembly. Through this approach, 191.5 Gb of high-quality 45–101 bp reads was generated, thus providing a 65.6-fold coverage of the caprine genome.

According to Dong et al. [103], about 42% of the goat genome contains repetitive elements, with many of them being ruminant-specific. Comparison with cattle revealed an expansion, in goats, of SINE-tRNA repeats as well as a contraction of SINE-BoVa elements. A consensus gene set of 22,175 protein-coding genes, 262 rRNA, 829 tRNA, and 1,010 small nuclear RNA genes was built, and evidences of
positive selection were found at several genes related with the immune response. Moreover, a total of 487 genes encoding microRNAs (miRNA) were detected, with six of them being goat-specific. A total of 157 miRNA genes were distributed in 44 clusters of variable size (2–46 miRNA genes), a similar proportion to that observed in humans. Structural variation of the goat genome has also been characterized with a comparative genomic hybridization array with ∼385,000 bovine probes [129]. A total of 127 copy number variation regions (CNVR), with an average size of 90 kb, were identified in Saanen, Camosciata delle Alpi, Girgentana, and Murciano-Granadina goats. Importantly, many of these CNVR also segregate in cattle suggesting ancient CNVR formation events in regions of genomic instability present in the ancestor of these two species [129].

Next generation sequencing has also been applied to characterize the transcriptome of goats [103, 130] as well as to generate large collections of SNPs mapping to transcripts [131]. Cashmere fine hair fiber, mainly produced in India and Mongolia, is derived from the hair secondary follicles that form part of the undercoat of Cashmere goats. Gene expression profiles of primary and secondary hair follicles of Cashmere goats have been compared in order to gain new insights into the genetic architecture of fiber quality [103]. The majority of expressed mRNAs corresponded to keratin and keratin-associated protein genes. Moreover, 28 downregulated and 23 upregulated mRNAs were identified in secondary versus primary follicles [103]. Amongst the differentially expressed genes, there were two keratin genes, 10 keratin-associated protein genes, fibroblast growth factor 21, asparagine synthetase, phosphoserine aminotransferase, and desmoglein 1. In another study, Geng et al. [130] compared skin mRNA expression at different hair follicle developmental stages (i.e., anagen, catagen, and telogen), thus identifying hundreds of differentially expressed genes. There has also been a sudden burst of studies analysing the expression profiles of miRNAs in the mammary gland [132–135], testis [136], and skin and hair follicles [137]. Interestingly, overexpression of caprine miR-103 in cultured mammary epithelial cells has been shown to have broad effects on lipid metabolism [135], stimulating the expression of genes related with lipogenesis, unsaturated fatty acid synthesis, triglyceride synthesis, cholesterol transport, fatty acid uptake, and fat droplet formation and downregulating those involved in lipid catabolism (lipolysis and fatty acid β-oxidation). In this context, it would be interesting to investigate if variability at miRNA genes and their binding sites could have important effects on milk composition and other traits of economic interest.

The design, by the International Goat Genome Consortium, of a 52K SNP chip for the high throughput genotyping of caprine DNA samples should be envisaged as a major scientific achievement in the field of goat genomics [138]. Polymorphism discovery was achieved by next-generation sequencing of dairy (16 individuals from the Alpine, Creole, and Saanen breeds sequenced in a HiSeq2000 platform with 13–26-fold coverage) and meat (pool 1: 20 Boer goats, pool 2: 20 Savanna goats and 24 Katjang goats, both sequenced in an Illumina Genome Analyzer IIx machine with 35-fold coverage) goat DNAs. Besides, reduced representation libraries for 17 Dutch Saanen goats were also sequenced with a low coverage. Altogether, this approach led to the identification of 11,924,638 variants including 1,229,120 indels and 10,695,518 SNPs. A total of 60,000 SNPs were selected and Illumina optimized assays for a total of 53,347 SNPs that were subsequently tested in a sample of 288 goats with a high success. This SNP chip has been already used by Kijas et al. [139] to carry out a genome-wide association study (GWAS) for polledness in Boer, Cashmere, and Rangeland goats. A highly significant GWAS signal was detected at chromosome 1, just at the genomic region reported by Pailhoux et al. [86] as containing the causal mutation for the goat polled intersex syndrome. Curiously, not all intersex goats were homozygous for the PIS deletion suggesting the existence of genetic heterogeneity for this reproductive pathology [139]. The 52K SNP chip has also been used to characterize the genetic diversity of Italian goat breeds [140], which showed a remarkable level of population structure, and it is the main tool employed in current international efforts to investigate goat genetic variation at a worldwide scale (Goat Adaptmap project, http://www.goatadaptmap.org/) as well as to develop genetic markers that can be used as selection criteria for phenotypes of economic interest (3SR project, http://www.3srbreeding.eu/ThenbspProject.aspx).

4. Final Remarks

In the last decades, the investigation of the genetic architecture of traits of economic interest in goats did not experience a substantial progress. Very few QTL studies were carried out and, consequently, there was a lack of positional information for genetic factors regulating phenotypic variation. A consequence of this state of things is that a very small number of causal mutations have been identified in goats. In contrast, many studies analysing physiological candidate genes and their association with meat and dairy traits have been carried out. However, in the absence of positional information and replication in diverse populations, the scientific and practical significance of these studies is quite limited and they cannot be used to establish gene-assisted selection schemes (with the exception of the CSNISI and PIS loci where consistent genetic models with causal relationships have been successfully established).

The recent construction of a 52K SNP chip [138] is expected to accelerate the speed at which discoveries are made, facilitating the performance of GWAS with an unprecedented resolution. Importantly, the 52K SNP chip produces data that can be easily standardized among labs at an affordable cost (at least if compared with microsatellites). In the near future, this should result in the detection, in a broad array of populations with diverse origins, of genomic regions with important effects on the phenotypic variance of quantitative and Mendelian traits and, eventually, in the identification of causal mutations that can be used as selection criteria. An additional benefit of the 52K SNP chip is that it can be also used to detect CNVR, although with limited resolution and some bias (SNP mapping to CNVR are underrepresented in chips because very often they are not...
in Hardy-Weinberg equilibrium). Copy number variation can have important phenotypic effects in domestic species [141], but so far there is only one report in goats and data were generated with a bovine 384K oligonucleotide array [129]. A future goal would be to increase the resolution of the caprine CNVR map and to investigate the association of structural variation with phenotypes. Finally, the invention of a 52K SNP chip also paves the way to apply genomic selection to goats, although the benefit-cost ratio of establishing such strategy should be carefully evaluated.

The high coverage sequencing of a goat genome [103] is another essential milestone that is expected to generate huge amounts of biological data and speed up research on themes of paramount importance, such as the physiological genomics of lactation, growth, and reproduction. The existence of a goat reference genome will also facilitate the monitoring of gene expression through RNA-seq in distinct goat tissues or experimental conditions, for example, goats differentially exposed to diets, thermal stress or pathogens, and so forth. The large-scale measurement of mRNA expression has been greatly facilitated by the development of next generation sequencing techniques, and it should play a crucial role in understanding how genetic polymorphisms and environmental factors modulate gene expression through complex biological networks that evolve in time and space. In this regard, genome sequencing and RNA-seq experiments have allowed to establish an initial catalog of the miRNA genes residing in the goat genome [103, 132–137], and in the next years we can anticipate that this list will be expanded to new miRNAs as well as to a plethora of regulatory RNAs, such as small interfering RNAs, Piwi-associated RNAs, and long noncoding RNAs, with significant effects on gene expression.

Massive parallel sequencing could also be applied to detect selective sweeps in the goat genome produced by artificial or natural selection. It should be taken into account, however, that processes as polygenic adaptation that can be fairly common in domestic species do not leave the classical genomic signatures observed for hard selective sweeps [142]. From a population genetics perspective, much information about the identification of goat domestication centers and genetic relationships amongst caprine breeds has been obtained from classical markers, as mitochondrial and microsatellite loci. However, the implementation of the 52K SNP chip and next generation sequencing techniques means that, in the coming years, powerful genome-wide approaches will be used to assess goat population structure at a worldwide scale as well as to investigate the influence of evolutionary forces on it (e.g., Nextgen project, http://nextgen.epfl.ch/). The participation of wild species other than the bezoar in the goat domestication process (either at primary or secondary domestication sites) could be also analysed by next generation sequencing. Another fundamental area of research would be to identify the genes that drove the behavioral and physiological changes leading to domestication. In this regard, the analysis of ancient DNA should provide new perspectives into this unresolved question. Finally, massive parallel sequencing could be also employed to characterize the diversity of highly endangered goat breeds and to devise strategies that ensure the conservation of such irreplaceable genetic resources.

**Conflict of Interests**

The author declares that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

Thanks are due to Drs. Juan Capote, Juan Manuel Serradilla, Baltasar Urrutia, and Juan Manuel Carrizosa for providing the goat pictures shown at Figure 1.

**References**


L. Ramunno, G. Cosenza, A. Rando et al., “Comparative
Y. Chilard, J. Rouel, and C. Leroux,” Goat’s αs1-casein in
M. Barbieri, E. Manfredi, J. M. Elsen et al., “Influence du
A. Zullo, C. M. A. Barone, L. Chianese, P. Colatruglio, M.
E. Manfredi, G. Ricordeau, M. E. Barbieri, Y. Amigues, and
P. Martin, M. Ollivier-Bousquet, and F. Grosclaude, “Genetic
G. Cosenza, R. Illario, A. Rando et al., “Characterization of the
M. H. Yahyaoui, A. Angiolillo, F. Pilla, A. Sanchez, and J. M.
O. C. Jann, E. Prinzenberg, G. Luikart, A. Caroli, and G. Erhardt,
M. Ambrosoli, L. Di Stasio, and P. Mazzocco, “Content of αs1-casein and coagulation properties in goat milk,” Journal of Dairy
M. Pazzola, M. L. Dettori, E. Pira, A. Noce, P. Paschino, and G. M. Vacc, “Effect of polymorphisms at the casein gene cluster...


[64] Q. J. Jin, J. J. Sun, X. T. Fang et al., "Molecular characterization and polymorphisms of the caprine Somatostatin (SST) and SST Receptor 1 (SSTR1) genes that are linked with growth traits," *Molecular Biology Reports*, vol. 38, no. 5, pp. 3129–3135, 2011.


137. C. Yuan, X. Wang, R. Geng, X. He, L. Qu, and Y. Chen, “Discovery of cashmere goat (Capra hircus) microRNAs in skin...


Submit your manuscripts at
http://www.hindawi.com