UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

DOTTORATO DI RICERCA IN SCIENZE VETERINARIE
XXXVI CICLO

TESI

STUDY OF GENES INVOLVED IN MILK PRODUCTION IN LIVESTOCK SPECIES

Tutor:
Prof.ssa Francesca Ciotola

Candidato:
Mariagiulia Pugliano

Coordinatore:
Prof. Paolo De Girolamo

INDEX

ABSTRACT

The main objective of this thesis was to genetically characterize populations of animals of zootechnical interest in central-southern Italy, focusing on genes associated with milk production and resistance to livestock diseases, in particular technical ones such as mastitis. Using the candidate gene approach, three native Lazio goat breeds (Capestrina, Grigia Ciociara and Bianca Monticellana) were analyzed at the calcium-sensitive casein loci (*CSN1S1*, *CSN2*, *CSN1S2*), the Bagnolese sheep at the *DIGAT1* locus, the *MBL2* gene in the Italian Mediterranean Buffalo and in the Comisana sheep , with the aim of characterizing them and subsequently associating them with production parameters and disease/health status. The choice of TGA (Autochthonous Genetic Types) and/or of breeds that contribute to the identity of a territory (Italian Mediterranean Buffalo), as well as the *loci* to be analyzed, is motivated by the intention to search for the peculiarities of the breeds that allow them to be distinguished from those with a high diffusion rate, to be able to plan their mating in order to avoid the loss of typical genetic variants, to define the productions to which they are more predisposed and to implement appropriate valorization plans.

The results of these studies have not only allowed us to present new, more accurate screening methods for certain *loci* (*CSN2* and *DGAT1*) but have also provided important information on the characteristics of the breeds and species analysed at these *loci*.

In particular, molecular analyses of calcium-sensitive caseins have revealed a discrete genetic variability for the *CSN1S1* locus in native goat populations of the Lazio region, in which four alleles (αs1-Cn A*, B*, F and N) have been identified, highlighting the almost complete absence of alleles associated with an intermediate or absent content of this protein in milk. When the three genetic types were grouped together, a high frequency of alleles associated with low (CSN1S1F) and high (CSN1S1A*, B*) αs1-casein content in milk was observed. Finally, the association with production parameters showed that animals with the A*A* genotype appear to be more productive than other genotypes in terms of fat and protein in milk, whereas animals with the FF genotype tend to produce milk with lower protein, casein and solids-non-fat content.

Molecular analysis of the *DGAT1* gene revealed a high frequency of the C allele (0.5655) at the g.5553 (C>T) locus, whereas the C allele predominated at the g.8539C>T locus. All animals tested for milk had a monomorphic genotype at g.8539C>T of the *DGAT1* gene, i.e. CC. Regarding the genotype at g.5553C>T, significant differences were observed between the genotypes in terms of percentage fat and protein yield. In particular, animals with the CT genotype had a higher percentage of fat produced per milking than those with the CC and TT genotypes ($p < 0.01$). Similar results were found for protein yield percentage, with CT individuals being more productive than CC individuals ($p < 0.01$).

The characterization of the *MBL2* gene in the Italian Mediterranean Buffalo showed that this gene is particularly polymorphic in this species, in fact 33 SNPs were found compared to the reference sequence, of which some (6689061 T>C, 6689062 G>A, 6689080 T>A, 6689100 G>A, 6690721 G>T, 6690757 C>A, 6690949 C>T, 6691003 T>C, 6691397 A>G) were statistically associated with SCC ($p<0.05$). Finally, regarding the *MBL2* gene in sheep, the characterization of exon 1 revealed the presence of three alleles: A1, C and D. The A1 variant is, in our opinion, a derivation of the A allele previously found in other sheep breeds. The A1 allele is responsible for an amino acid substitution in the corresponding translated sequence. Furthermore, variant A1 was the most frequent in the analyzed population (0.416). Interestingly, the A1C genotype was found to be present only in mastitis-free (healthy) animals, suggesting a possible relation between this genotype and the absence of mastitis; this hypothesis could explain the statistical significance observed in the analysis of the association between the A1C genotype and low SCC compared to the A1A1 and DD genotypes.

In this context, these studies make an important contribution to the genetic understanding of Italian livestock breeds and to the promotion of sustainable livestock production, with practical implications for genetic selection and livestock health management.

INTRODUCTION

The selection of ruminants for breeding purposes is a multifaceted process that involves balancing genetic improvement, adaptability, economic efficiency, and environmental issues. Ruminants play a key role in global agriculture, contributing significantly to meat and milk production.

Animal breeding is a process in which phenotypic modifications implemented through genetic selection are often necessary in order to improve traits of economic-productive interest such as milk and meat production and disease resistance. This process, also known as genetic improvement (GI), aims to achieve sustainable improvement of desired traits in farm animals.

In order to achieve genetic improvement within a breed, it is essential to develop and implement a selection plan aimed at spreading in the population under selection only the most desirable alleles in subsequent generations. Favorable alleles are defined on the basis of pre-established selection objectives which are based on the species/breeds, their productive aptitude and farm goals.

To achieve genetic progress, it is essential to assess and select which and how many animals should be recruited to procreate the next generation. The aim is to pass on the desired genetic traits to the next generation, gradually increasing the overall performance of the population.

Although the objectives of selection are different for each species / breed under selective pressure, the reason that drives breeders and associations in the sector to perpetuate selection is basically the same: to make the breeding of that species economically advantageous, in order to guarantee a good profit for the individual breeder and the community. This objective was initially pursued by focusing exclusively on improving the qualitative and quantitative characteristics of production (milk production, fat production, protein production, improving the percentage of fat and protein which, in the case of species whose milk is processed, can be translated into a better cheese-making yield, etc.). However, it was later realized that the exclusive focus on productivity had a negative impact on animal health and welfare. Consequently, the course was reversed, and emphasis was placed on traits such as disease resistance, fertility, longevity, etc. This not only improved animal welfare, but also reduces the management costs of the individual animal and the herd, thus benefiting the production system. In essence, all this has so far resulted in improved primary (increased profits) and secondary (reduced costs) aspects of production.

The main tool for optimal management of selection plans and thus of the genetic improvement of livestock is the herd book. This is the official list of animals belonging to a breed whose parentage and productive potential are known, as they are subject to functional controls. Breed associations are responsible for managing the herd books, delegating operational activities to territorial bodies such as provincial associations.

Initially, herd books were open, which meant that animals could be registered without knowing their pedigree, provided they had the phenotypic traits typical of the breed. Later they became closed, restricting entry to animals whose pedigree was known to at least the third or fifth generation and which were also registered in the herd book.

Each animal recorded in the herd book is provided with a document called 'pedigree' which certifies its identity. This document contains information such as name, registration number, date of birth, owner's name and genealogical chart. Throughout the life of the registered animal, the herd book records a wide range of information useful for selection purposes, including productive, reproductive and morpho-functional aspects.

Today's selection targets are very different from those of 10, 20 or 30 years ago. Given the challenges posed by climate change and unpredictable environmental conditions, prioritizing adaptability and resilience in ruminant selection is of paramount importance. Breeding initiatives aimed at improving traits such as heat tolerance, disease resistance and efficient use of resources are essential to ensure the ability of the herd to thrive in a range of conditions. It is therefore easy to see that current selection objectives cannot only include production issues but must also address overall aspects of animal resistance and resilience.

The evolving targets in the livestock farms are reflected in the changes of breeding objectives applied to dairy animals over time, with a particular emphasis on the bovine species. A detailed investigation conducted by Miglior et al. (2017) revealed a significant evolution in the focus of genetic selection in dairy cattle populations, spanning from the early 20th century to the present day.

Initially, in the early decades of the last century, the predominant focus was on increasing milk production. Breeders directed their efforts toward traits closely associated with the quantity of milk produced, clearly defining selection objectives.

However, over time, a notable shift in this approach occurred. Towards the end of the century, there was a significant move towards a more balanced breeding goal. Consideration began to extend beyond just milk quantity, incorporating traits such as milk quality, longevity, fertility, calving ease, health, and workability. This evolution reflects a growing awareness of societal influence and an enhanced understanding of animal physiology.

Another pivotal stage was the broadening of selection objectives to include economically relevant traits with low heritability. This expansion was facilitated by the awareness that such traits could be genetically improved, aided by modern technologies providing more extensive data. Consequently, new phenotypes were introduced into breeding objectives, with a specific emphasis on fitness, health, welfare, milk quality, and environmental sustainability.

Another important moment was the recognition that animal selection for few specific traits, such as conformation and production, had indirect negative effects, on others, such as health and fertility. To counter these negative trends, new selection indices were developed, balancing health and fertility more equitably and contributing to genetic progress on all fronts.

In recent years, a further step forward has been taken with the use of Milk Infrared Spectral (MIR) data. This technology has opened new opportunities for selection, allowing for more precise prediction of fine milk components and coagulation properties. This represents a shift towards a more detailed assessment of milk properties and quality, with a greater focus on production intended for human consumption and cheese production.

In summary, the evolution of the focus of genetic improvement in dairy cattle has transitioned from a narrow emphasis on milk production to a more comprehensive and balanced breeding goal, embracing a wide range of economically relevant traits, including health, fertility, and milk quality.

Considering dairy industry, the main focal points for animal selection are production traits in terms of quantity associated with quality aspects, which include milk yield, protein and fat percentages, absolute protein and fat content and more. At the same time, secondary traits related to disease resistance and resilience are important. The goal of improving understanding of the genetic basis regulating primary and secondary production traits has driven several studies over the years. These investigations aim to unravel the complexities of the genetic basis of milk quality attributes as well as disease resistance and resilience traits in livestock.

Studies focusing on the genetic bases of milk quality have uncovered lot of genes involved in this trait, among them there are *CSN1S1, CSN1S2, CSN2, CSN3* (encoding for αs1, αs2, β, and κ casein respectively), *LALBA* (alpha-lactalbumin), *BLG* (beta-lactoglobulin), *LTF* (lactoferrin), *DGAT1* (diacylglycerol O-acyltransferase 1), *OXT* (oxytocin), *STAT1*(transcription signal transducer and activator 1), *STAT5*(transcription signal transducer and activator 5A), *SCD* (stearoyl-CoA desaturase), *ABCG2*(ATP-binding cassette subfamily G member) (Singh et al., 2014, Selvaggi et al., 2014a, Selvaggi et al., 2014b, Ibeagha-Awemu et al., 2008).These genes are crucial in shaping the characteristics of milk, which contributes significantly to our understanding of genetic factors influencing milk quality.

As regard resistance and resilience traits the genes that, according to the most important studies, seem to be mainly involved are *MBL2, LTF, SLC11A1, TLR2, TLR4.* These genes play key roles in the immune response and defense mechanisms of mammals. In particular, *mannose binding lectin 2 (MBL2)* is associated with the innate immune system, *lactoferrin (LTF)* is known for its antimicrobial properties, *Solute Linked Carrier 11A1 (SLC11A1)* is involved in regulating intracellular survival of microbes, *toll-like receptor 2 (TLR2)* and *toll-like receptor 4 (TLR4)* are key players in recognizing

and responding to pathogens (Singh et al., 2014). The latter gene has also recently been correlated with milk quality by Wang et al. (2018); indeed, statistically significant effects ($P < 0.05$) were shown for SNP -226 G > C, which influences milk yield on the day of testing (TDMY), fat content (FC), protein content (PC), total solids (TS) and milk urea nitrogen (MUN).

Among the genes mentioned above, the focus of this study was on *CSN1S1*, *CSN1S2*, *CSN2*, *DGAT1* and *MBL2*.

GENES RELATED TO MILK PRODUCTION

Ongoing research in livestock genomics has significantly advanced our understanding of the genetic basis of key animal traits, particularly in the context of dairy production.

Regarding the buffalo species, considering only studies using the candidate gene approach, a modest number of genes have been identified, i.e. 19 candidate genes (Table 1), carrying 47 mutations linked to different milk production traits in different buffalo breeds. Among these mutations, 19 were located within coding regions, with 13 causing amino acid substitutions. The remaining 24 mutations were scattered across intronic regions, while four were in promoter regions (Du et al., 2019). With the advancement of novel techniques for genomic studies in livestock, an expanding list of genes has emerged, showing potential associations with milk production and disease resistance. In buffalo species, employing the 90k SNPchip (AXIOM) and the Genome-Wide Association Study (GWAS) approach, Du et al (2019) uncovering 499 genes potentially linked to production traits. Notably, among these, *CTNND2* (catenin delta 2), *APOB* (apolipoprotein B), *FHIT* (fragile histidine triad), and *ESRRG* (estrogen-related receptor gamma) stood out, having been previously identified in other studies. In particular *APOB* has shown interaction relationships with candidate genes such as *LEP* (leptin), *SREBF1* (sterol regulatory element-binding transcription factor 1) and *TG* (thyroglobulin), thus it could be an interesting gene for further study in buffalo production genetics.

Moreover, through a comprehensive bioinformatic analyses of the genome a total of 517 genes closely linked to buffalo milk production traits have been identified (Du et al., 2020). Subsequent analyses delved deeper into these genes, unraveling the intricate molecular mechanisms governing protein-protein interactions pivotal to milk production. The bioinformatics methodology employed in this research establishes the scientific bases, enhancing our comprehension of the genetic and molecular intricacies underpinning buffalo milk production, and offering promising avenues for selective improvement.

Extensive GWAS analyses have been conducted in cattle as well. These analyses led to the annotation of several genes associated with both milk production and disease resistance. For instance, Ilie and colleagues (2021) conducted a GWAS analysis on two Romanian cattle breeds and identified 41 SNPs associated with Somatic Cell Score (SCS). Notably, four of these SNPs were in close proximity to genes *HERC3*, *LUZP2*, *AKAP8*, and *MEGF10*, as identified in other studies, indicating them as potential candidates for SCS.

Furthermore, Iung et al. (2019) unearthed potential novel genes linked to various traits, including tissue damage and repair of the mammary gland (*COL18A1*), immune response (*LTTC19*), glucose homeostasis (*SLC37A1*), synthesis of unsaturated fatty acids (*LTBP1*), and sugar transport (*SLC37A1*

and *MFSD4A*). These genes were found to be also linked with traits such as milk yield, somatic cell score, fat percentage, and fatty acid composition.

In sheep and goats, GWAS studies have led both to the confirmation of the economic importance of some widely studied genes (*CSN1S1*, *CSN1S2*, *CSN2*, *CSN3*, *DGAT1*) (Massender et al., 2023) and to the annotation of new potential candidate genes for both milk quality parameters and conformation traits. In particular, the studies conducted by Argyriadou et al. (2023) in Chios and Frizarta sheep identified nineteen candidate genes (*POT1*, *TMEM229*, *NTAQ1*, *ZHX1*, *ZHX2*, *LOC101109545*, *HAS2*, *DERL1*, *FAM83A*, *ATAD2*, *RBP7*, *FSTL1*, *CD80*, *HCLS1*, *GSK3B*, *GRID2*, *FAIM*, *CEP70* and *GRIP1*) potentially associated with animal resilience traits including somatic cell count (SCC), lactation persistence (LP) and body condition score (BCS). Sutera et al. (2021) added further detail to the understanding of somatic cell count (SCC) genetics by identifying 34 genes potentially associated with SCC traits in Valle del Belice sheep. The most interesting among these genes was the Stress Associated Endoplasmic Reticulum Protein 1 (*SERP1*) gene, suggesting that it may play a key role in regulating somatic cell count in sheep milk.

Studies by Selionova et al. (2024) and Massender et al. (2023) provided important insights into the genetics of milk composition in goats, identifying numerous genetic variants and candidate genes associated with different milk quality and conformation parameters.

Selionova et al. (2024) identified 66 candidate genes significantly associated with fatty acid composition in goat milk. Of these, 12 had a polygenic effect: most were simultaneously associated with dry matter content and fatty acids (*METTL*, *SLC1A8*, *PHACTR1*, *FMO2*, *ECI1*, *PGP*, *ABCA3*, *AMDHD2*).

On the other hand, Massender et al. (2023) identified 271 candidate genes for milk quality and conformation traits in the Saanen and Alpine Canadian breeds. Interestingly, common candidate genes for milk production traits were identified in n these two breeds (*ADAMTS3*, *CLOCK*, *CSN1S1*, *CSN1S2*, *CSN2*, *CSN3*, *DCK*, *LOC102178810*, *LOC102179276*, *LOC106502214*, *MOB1B*, *SHROOM3*, *SLC4A4*, *TRNAC-ACA* and *TRNAC-GCA*). Furthermore, in the same study, the authors identified SNPs located in the casein cluster and in the *DGAT1* gene region associated with milk quality traits such as protein and fat content, confirming the economic importance of these genes. According to the same study (Massender et al., 2023), other genes of particular interest are *RPL8* (encoding 60S ribosomal protein L8), that may play an important role in fat synthesis in cattle, *DCK* (encoding deoxycytidine kinase) and *MOB1B* (encoding MOB 1B kinase activator), both genes have been associated with udder health traits in dairy cattle.

These studies therefore highlight the complexity of the genetics of milk production and conformation in dairy animals and provide new perspectives for improving genetic selection and herd management to optimize milk production and animal welfare.

Table 1. Genes identified using candidate gene approach in buffalo (Du et al., 2019).

Gene name (accession no.)	**Buffalo breeds /origin (sample size)**	**Mutations (accession no.)**	**Amino acid substitution**	**Associated milk production traits**	**References**
ADRA1A (NW_005784806)	Murrah Buffaloes (220)	Exon 2: g.170C>T	No	Protein percentage	Araújo et *al*. (2015)
A2M (NW_005785373)	Murrah Buffaloes (136)	Exon 29: g.241A>G (AC_000162.1)	Il suo ad Arg	Fat yield, fat percentage, protein percentage	Freitas et *al*. (2016a)
BTN1A1 (NW_005784078)	Murrah Buffaloes (55)	Intron 1: c.184C>T>G		Milk yield	Kale et *al*. (2013,2014)
CSN1S1 (NW_005785429)	Italian Mediterranean Buffaloes (175)	Exon 17: c.628C>T (AY948385)	Ser178 Leu	Protein percentage	Cosenza et *al*. (2014)
DGAT1 (NW_005785136)	Murrah Buffaloes (196)	Exon 17: g.11785T>C (AY032689)	Ala484 Val	Fat percentage, Protein percentage	Freitas et *al*. (2016b)
	DeHong Buffaloes (81)	Intron 8: 14 C>T; 35 G>T (AY999090)		Fat percentage	Fan *et al*. (2011)
GHRL (NW_005785786)	Brasilian buffaloes(240)	Intron 3: g.778C>T; g.960G>A		Fat yield	Gil *et al*. (2013)
		Intron 3: g.905T>C		Fat yield, fat percentage, protein percentage	
INSIG2 (NW_005785215)	Murrah Buffaloes (55)	Intron 1: g.621272A>G; g.621364A>C Intron 2: g.632543G>A; g.632684C>T		Milk yield	
	Murrah × GuangXi (80) e Nili-Ravi × GuangXi (66) Buffaloes	Intron 1: g.621272A>G; g.621364A>C Intron 2: g.632543G>A; g.632684C>T KF975668.1		Milk yield, protein percentage	
LALBA (NW_005784627)	Bhadawari Buffaloes (48)	Exon1: g.35T>G;	Ile a Ser	Milk yield	Dayal et *al*. (2006)
		Exon1: g.95A>G;	Lys to Arg		
		Exon1: g.103A>G; g.104A>C	Lys ad Ala		
		Exon1: g.115G>T	Gly a Cys		

LEP (NW_005782251)	Mehsana Buffaloes (154)	Intron1(g.98C>T, g.111G>A, g.172G>A, g.209T>C>A, g.266G>A) (AJ512638)		Fat percentage	Tanpure et *al*. (2012)
	Egyptian buffalo (300)	Intron2: g.83A>G		Milk yield, fat percentage	Nasr et *al*. (2016)
MC4R (AC_000181)	Chinese buffalos (332)	Exon region: g.1104C>T (KJ139982.1)	No	Milk yield, fat percentage, protein percentage	Deng et *al*. (2016b)
MTNR1A (NW_005785680)	Brasilian buffaloes (200)	Exon2: g.72T>C (JN689386)	No	Protein percentage	Zetouni et *al*. (2014)
OXT (NW_005782921)	Italian Mediterranean Buffaloes (163)	Exon2: g.1627G>T (AM234538)	Arg a Leu	Milk yield	Pauciullo et *al*. (2012a)
PRL (NW_005785510)	Italian Mediterranean Buffaloes (465)	Intron1: g.2609C>T		Milk yield, protein percentage	Li *et al*. (2017)
		Exon2: g.2836C>T (102412882)	Arg a Cys	Fat percentage	
SCD (NW_005785261)	Italian Mediterranean Buffaloes (169)	Promoter region: g.133A>C		Milk yield	Pauciullo et *al*. (2012b)
	Nili-Ravi Buffaloes (346)	Exon4: g.8453A>T (AY241932)	Lys158Ile	Fat percentage	Maryam et *al*. (2016)
SPP1 (NW_005782435)	Nili-Ravi Buffaloes (50)	Exon4: g.38329758T>C (NC_007304.5)	Val127 Ala	Protein percentage	Manzoor et *al*. (2017)
SREBF1 (NW_005770853)	Murrah × GuangXi e Nili-Ravi × GuangXi Buffaloes (384)	Exon2: g.8402C>T	No	Milk yield	Deng et *al*. (2017)
		Intron18 g.15369G>T (KX951420.1)		Milk yield, fat percentage	
STAT1 (NW_005783879)	Murrah × GuangXi(112) e Nili-Ravi × GuangXi (80) Buffaloes	Exon10: 15642G>T	Leu280 Arg	Milk yield, fat percentage, protein percentage	Deng et *al*. (2016c)
		Intron2: 2338T>C; Intron4: 5558G>T		Milk yield	
		Intron2: 1079T>C; 2965A>C Intron3: 3666G>A; 4007T>C (KJ139981.1)		Protein percentage	

STAT5A (NW_005784710)	Binlangjiang buffaloes (83)	Exon8: 924C>T	No	Milk yield	Min *et al*. (2013)
		Exon8: 975C>T (NW 003104496.1)	No	Fat percentage	
TG (NW_005785850)	Mehsana (136) e Nili-Ravi (82) buffaloes	promoter region: 33C>T; 176G>T; 221C>T		Fat percentage	Dubey et *al*. (2015)

Casein genes

Caseins are the main milk proteins, representing approximately 80% of the milk's protein nitrogen (Lajnaf, et al., 2022). They are conjugated proteins, i.e. they are bound to other molecules, and, in particular, to phosphorus in the form of esterified phosphoric acid. These phosphoric groups are important both for the protein's structure and its function.

For infants, caseins represent the primary source of amino acids and play a key role in transporting phosphorus and calcium in quantities that are sufficient for bone tissue development. Specifically, caseins from the αs2 and αs1 groups, being the richest in phosphorus, determine the ability of micelles to transport calcium and phosphorus in colloidal form (Ramunno et al., 2007).

As known, in ruminant milk, there are four casein fractions (αs1, β, αs2, and k), encoded by four autosomal genes (*CSN1S1*, *CSN2*, *CSN1S2*, and *CSN3*, respectively). Their structure allows their association to form micelles, which, in turn, have a hydrophobic fraction consisting of the central core containing calcium-sensitive caseins (αs1, αs2, and β), and an external hydrophilic fraction primarily composed of k-casein, which stabilizes the micelle.

The three genes responsible for encoding calcium-sensitive caseins (αs1, αs2, and β), exhibit shared regulatory motifs in their proximal 5' flanking region (Groenen and Van der Poel, 1994).

The caseins comprising the hydrophobic fraction are held together through hydrogen bonds and bonds that also involve calcium ions. To keep the casein micelles in dispersion, it is necessary, for the milk, to have an appropriate pH, intact k-casein on the outside of the micelle, and a proper balance of calcium and phosphorus salts in the milk. If only one of these conditions are not met, the hydration state is no longer sufficient, and the micelles interact by flocculating or coagulating.

The polyphosphorylation sites are responsible for the calcium sensitivity of αs1, αs2, and β caseins, they constitute the anchoring points for phosphocalcic bridges, connecting the submicelles to each other. Calcium-sensitive caseins are currently the most studied, particularly in the caprine species, where significant genetic polymorphism is often responsible for variations in protein quantity and structure, correlating with changes in milk production and technological characteristics (Ramunno et al., 2007).

As reported by Selvaggi et al., (2014a,b) the four calcium-sensitive caseins exhibit similar molecular weights (around 24 kDa), promoter regions, leader peptide sequences, and locations of the major phosphorylation site.

A phylogenetic analysis conducted by Hassanin and colleagues (2022) on the sequences of the four casein genes of several ruminants, including the camel, showed that sheep, goats, cattle, and buffaloes can be divided into two distinct clusters. The first one, comprising sheep and goats, exhibits a high

similarity (98.06%) in caseins sequence, the other one, including cattle and water buffalo, shows a similarity of 96.25% (Fig. 1).

These data support the hypothesis of a common evolutionary origin for these genes, arising from duplications and modifications of a unique ancestral gene.

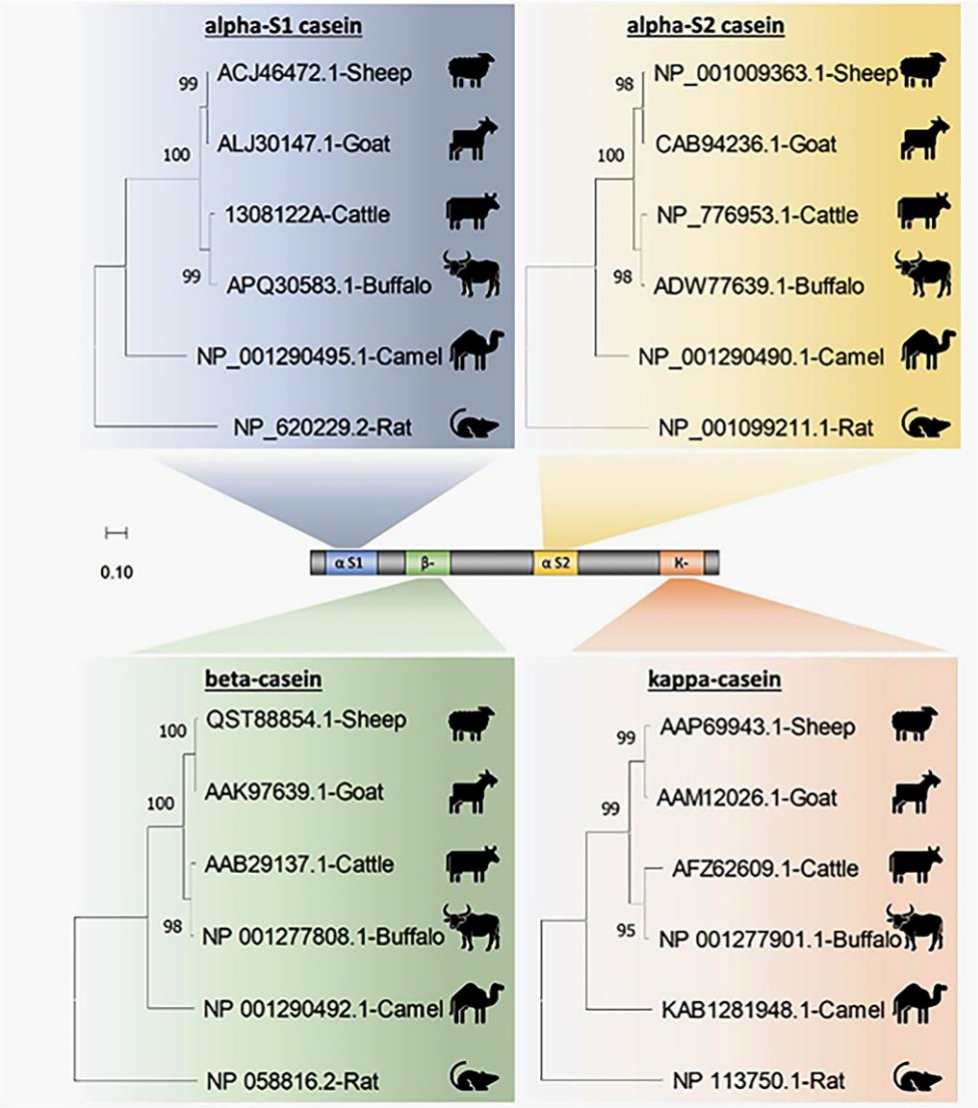

Fig.1. Phylogenetic analysis of casein subunits from different domestic animals. A neighbor-joining tree representing the phylogenetic relationship of sheep, goats, cattle, Arabian camels, and water buffalos was derived from the alignment of alpha-S1-, alpha-S2-, beta-, and kappa-casein amino acid sequences. The tree was rooted in *Rattus norvegicus*. Bootstrap support values are indicated on the nodes. Bootstrap values were calculated from 1,000 replications (Hassanin et al., 2022).

In all dairy livestock, the genes encoding caseins exhibit a distinctive arrangement, forming clusters on a specific segment of the chromosome spanning approximately 250 kb of DNA. This pattern is consistent in sheep, goats, and cattle, in which these genes are located on chromosome 6. Conversely, in buffaloes, this genomic region is identified on chromosome 7 (Fig. 2).

Another noteworthy aspect involves the close proximity of the *CSN1S1* and *CSN2* genes in all ruminant species. Their distance is remarkably short, merely 12 kb, and they are transcribed in a convergent manner (Leroux et al., 1992). This phenomenon is uniformly observed across various ruminant species (sheep, goat, cow and buffalo) suggesting a common genetic pattern. This arrangement is not limited to ruminants but is also observed in other mammalian families, including rodents, equids, camelids, and primates, including humans. The proximity and convergent orientation of these genes may underscore their functional relationship and the evolutionary significance of this specific genetic configuration across different species.

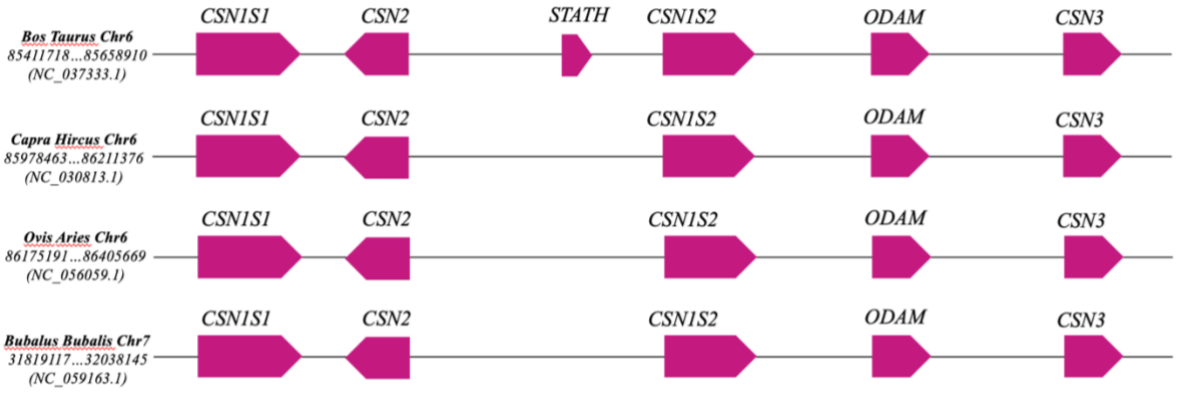

Fig. 2. Structural organization of the genes encoding the four milk caseins in ruminants

CSN1S1

αs1-CN is the main casein component of bovine milk, which can be up to 40% (Farrell et al., 2004), followed by buffalo milk with an average of 32% (Bonfatti et al., 2012b), while in goat milk a high variability in the amount (from 0 to 25%) and composition of this casein fraction is observed, depending on the presence of polymorphisms in the coding and regulatory sequences of the gene (Torres-Vazquez et al., 2008). With regard to the protein fractions in sheep's milk, it is the milk that contains the most caseins (CN) (76-83% of total protein), of which αs1-CN accounts for 6.66% (Anusha Siddiqui et al., 2024, Selvaggi et al., 2014b).

In general, within species, multiple variants of αs1-CN have been identified. In cattle, there are nine known protein variants of αs1-CN, classified as follows: A, B, C, D, E, F, G, H, and I (Caroli et al., 2008). In 2013, Gallinat et al. discovered a new genetic variant due to a nucleotide substitution in position 543 of exon 17 (c.543G > T) of the *CSN1S1* gene. This modification results in a change in amino acid composition, with valine (Val) being replaced by phenylalanine (Phe) at position 182 (p.Val182→Phe). This polymorphism has been identified in the bovine breeds Sarabi (*Bos Taurus*) and Girn (*Bos Indicus*). What makes this discovery particularly interesting is the fact that, up to that

moment, this specific genetic variant and the predicted association with the amino acid exchange had not been detected. The researchers named this new variant as *CSN1S1**J to classify and distinguish it from other variant of the *CSN1S1* gene.

Each variant differs from the others for changes in the amino acid sequence, with the exception of the G variant, which differs from the B variant only in terms of lower expression of the corresponding protein (Caroli et al., 2009; Rando et al., 1998).

Among these variants, B is the most frequent in *Bos taurus*, along with C, which is also widespread in *Bos indicus*. The C variant differs from B by the amino acid change Glu to Gly at position 192 (p.Glu192→Gly) of the mature protein (Farrell et al 2004; Roin et al., 2022).

In sheep, Selvaggi et al. (2014b) reported eight variants of CSN1S1, named A, B, C, D, E, F, H, and I, through protein electrophoresis techniques. The variations in amino acid substitutions and phosphorylation levels among these variants have been shown to impact both milk composition and its technological properties and in particular cheese yield. Variant D, commonly referred to as the Welsh variant, earned its name because it was initially discovered by King in sheep inhabiting the mountains of Wales in 1966. Currently, this variant is regarded as the least favorable from a production standpoint due to its observed adverse effects on milk composition and on the quality of the processed product. This can be attributed to the presence of a p.Ser68→Asn substitution, leading to the loss of two phosphorylation sites and making it the least phosphorylated variant (Selvaggi et al., 2014b; Ferranti et al., 1995).

Among all the ruminants of zootechnical interest, goats exhibit the highest variability at the *CSN1S1* locus, with a remarkable 17 allelic variants associated with 4 levels of synthesis of the corresponding protein, αs1-CN (Ramunno et al., 2008; Torres-Vazquez et al., 2008; Mangia et al., 2019). This unquestionably represents the most significant casein for the caprine species.

Based on the quantity of αs1-CN in the milk, it is possible to distinguish alleles with different production capacities. Strong alleles (A, B1, B2, B3, B4, C, H, L, and M) exhibit a production of approximately 3.5 g/l; intermediate alleles (I and E) have a production of around 1.1 g/l; weak alleles (D, F, and G) show a production of about 0.45 g/l; finally, null alleles (01, 02, and N) are responsible for the absence of casein in the milk (Jansa-Perez et al., 1994; Ramunno et al., 2007; Ramunno et al., 2008; Mangia et al., 2019; Cosenza et al., 2003) (Fig. 3).

Most alleles, especially those associated with a normal quantity of casein, originate from single nucleotide substitutions (Cosenza et al., 2008). However, some allelic variants result from insertions (alleles E, 02), deletions (alleles 01 and N), or both (allele F) (Jansa-Perez et al., 1994; Leroux et al., 1992; Ramunno et al., 2005; Cosenza et al., 2003; Martin et al., 1999).

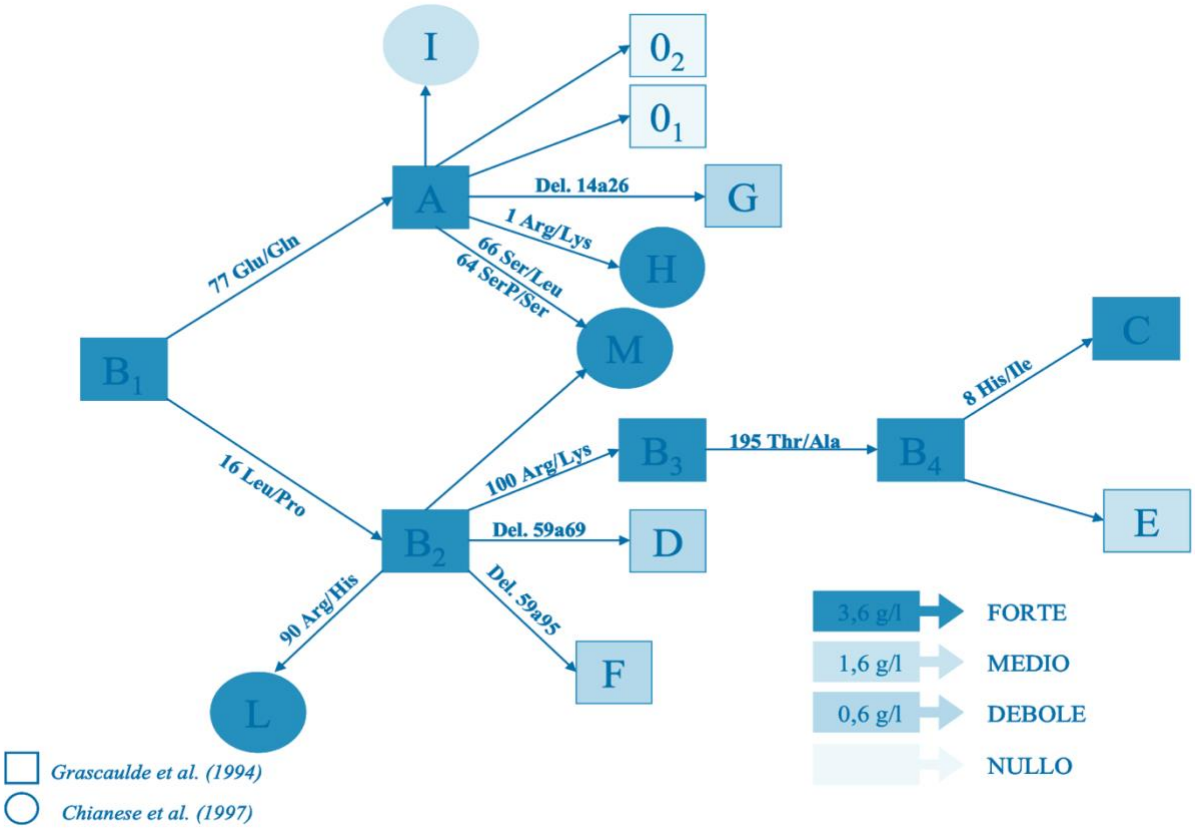

Fig. 3. Phylogenetic tree of α_{s1}-CN revised by Chianese et al. (1997)

In water buffalo (*Bubalus bubalis*), there have been few studies conducted so far. Indeed, three allelic variants (A, B and B^{RV}) of as1 casein have been identified in the Italian Mediterranean buffalo breed, in the Romanian buffalo breed and in Binglangjiang buffalo. These findings are documented in several studies (Chianese et al., 2009, Balteanu et al., 2008, Cosenza et al., 2015, Bonfatti et al., 2012a, Bonfatti et al., 2012b, Zhang et al., 2021). The single nucleotide polymorphism (SNP) responsible for the amino acid change is located at position 89 of the 17th exon (c.628C>T), resulting in the amino acid substitution p.Ser178→Leu, for allele B and A, respectively (Cosenza et al., 2015). A significant correlation has been observed between the c.628C>T SNP and the protein percentage in the milk. Specifically, the CC genotype exhibits on average approximately 0.04% more protein content in milk than the CT and TT genotypes (Cosenza et al., 2015). According to Bonfatti et al 2012b, genotypes with *CSN1S1 A* were associated with an increase proportion of αS1-CN in total casein (TCN).

Zicarelli et al. (2020) confirms what was stated by Cosenza and colleagues in 2015, namely that the milk of the Italian Mediterranean buffalo carrying the αS1-CN B variant resulted in a higher curd yield. Additionally, the differences in the raw protein percentage among the αS1-CN genotypes were

significant ($P < 0.01$), with the αS1-CN B variant specifically associated with a higher quantity of raw protein. These results suggest that even in the Italian Mediterranean buffalo, as in the other main dairy ruminant species, the genetic variant influences the protein composition of milk, with potential implications for milk production and quality.

Zhang et al. (2021) describes 8 haplotypes for this gene in Buffalo named as A, B, B', B'', C, D, E ed F (Fig. 4). According to his study B variant is the ancestral one from which B' and B" derived following a synonymous nucleotide change that unaffected ammino acid sequence. It has been suggested that C variant derived from B' due to c.620G>A substitution in *CSN1S1* gene. This SNP is reflected in the polypeptide chain with amino acid substitution p.Gly192→Glu (Cosenza et al., 2015; Sukla et al., 2007)

CSN1S1 A variant would originate from B for the SNP c.578T>C that causes the amino acid change in the polypeptidic chain p.Ser178→Leu (Cosenza et al., 2015; Bonfatti et al., 2012a) .

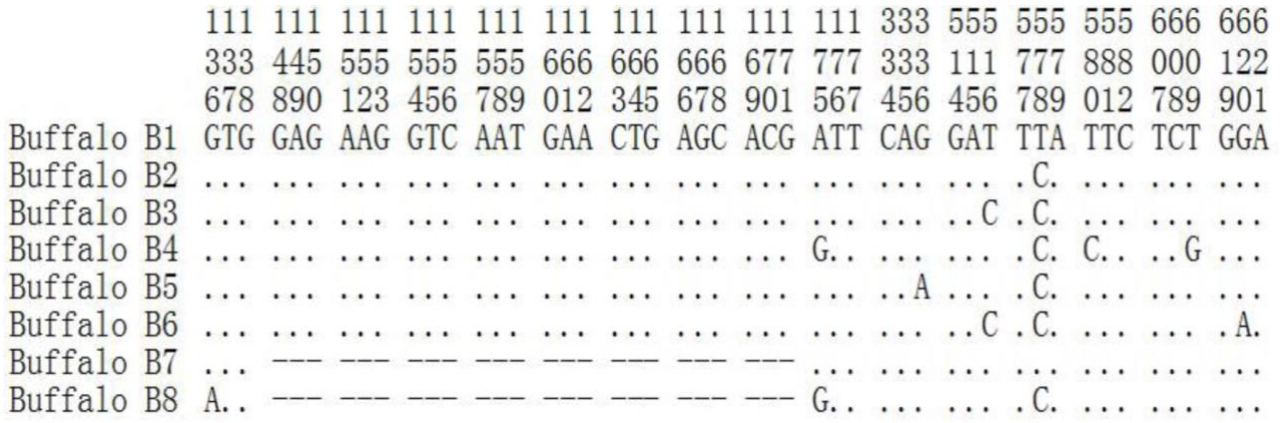

Fig. 4. Graphic representation of nucleotide substitutions along the *CSN1S1* gene in buffalo (Zhang et al., 2021).

The SNP in the intronic region g.472G>C of the *CSN1S1 A* allele gave rise to *CSN1S1 E* allele. This sequence change causes the inactivation of the splicing site and the skipping of exon 6 in the mRNA thus producing a protein missing 8 ammino acids (35-42). D and F variants share the SNP c.175A>G of exon 7, but they differ because the variant D has two SNP: the first one occurring at position c.580 is a transition T>C that gives rise to the protein ammino acid change p.Phe179→Leu and the second is a synonymous transversion c.609T>G in ex17 (Ser188→Ser). Moreover, F allele, that has sometimes been named B^{RV}, has the same deletion of E allele and, with the addition of the SNP c.136G>A that causes the ammino acid change Val31→Met, where Met is in the F allele (Cosenza et al., 2015). Finally, according to Zhang et al., (2021), D variant is only of the Swamp Buffalo while

the ancestral B variant is only in river buffalo. All protein variants are graphically summarized in Fig. 5

Variants (haplotype)	Positon and amino acid in the mature peptide								
	31	35-42	44	97	157	178	179	188	192
Buffalo A (B1)	Val GTG		Ile ATT	Gln CAG	Asp GAT	Leu TTA	Phe TTC	Ser TCT	Gly GGA
Buffalo B (B2)						Ser TCA			
Buffalo B'(B3)					Asp GAC	Ser TCA			
Buffalo B"(B5)				Gln CAA		Ser TCA			
Buffalo C (B6)					Asp GAC	Ser TCA			Glu GAA
Buffalo D (B4)			Val GTT			Ser TCA	Leu CTC	Ser TCG	
Buffalo E (B7)		deleted				Leu TTA			
Buffalo F (B8)	Met ATG	deleted	Val GTT			Ser TCA			

Fig. 5. Amino acid positions and differences in the variants of buffalo αS1-CN (Zhang et al., 2021).

CSN2

The *CSN2* gene encodes for β-casein, described as "calcium-sensitive" due to its precipitation in the presence of low concentrations of Ca cation. β-CN, after αs1-casein, is the protein component proportionally most abundant in the milk of ruminants. The β-CN family constitutes up to 45% of the casein in bovine and ovine milk (Farrell et al., 2004; Selvaggi et al., 2014b), while in goats, it is generally present in quantities around 6 g/l (Ramunno et al., 2007), representing the most predominant fraction. Similarly, it has been demonstrated that β-CN constitutes approximately 53% of the entire casein fraction for Italian Mediterranean buffalo (Cosenza et al., 2011), followed by the αs1 fraction. One of the key findings which prompted studies on this protein and its coding gene was its intense impact on the variation of RCT and K20, as highlighted by Bonfatti et al. in 2013.

The gene that encodes for β-CN is *CSN2*, which in ruminants spans approximately 9-10 kb (Cosenza et al., 2009).

At least twelve bovine variants for β-CN, namely A1, A2, A3, B, C, D, E, F, G, H1, H2, and I, have been reported (Farrell et al., 2004). Additionally, Gallinat et al. (2013) discovered 3 further non-synonymous variants named J, K and L according to the alphabetical order of casein variants. Using the variant A1 as reference, these three new variants are respectively characterized by c.103G>A, c.580C>G (in addition to c.245A > C) and c.635T>C resulting in Glu35→Lys, Pro194→Ala and Val212→Ala respectively. Additionally, a variant named A4 have been identified in Korean native cattle but it hasn't been characterized yet (Kaminski et al., 2007)

Ramesha et al. in 2016 explored genetic variants (A1, A2, A3, and B) in Bos *Taurus* and *Bos Indicus* populations, focusing on the amino acid differences encoded in exon 7. A1 and A2 were the most common variants, differing in the amino acid substitution p.His67→Pro for allele A1 and A2, respectively. Histidine found in A1 milk is a weak bond that is easily broken to release the bioactive peptide b-casomorphin 7, while proline found in A2 milk is a strong bond that does not break during digestion. The A2 variant is considered the oldest and the original variant from which all other variants are derived (Fig. 6) The most common genetic variants are A1 and A2, while variants B and I are usually identified at lower frequencies (Kaminski et al., 2007). A more recent analysis by Sebastiani et al. (2020) through a thorough literature review confirms that variant I is generally uncommon, while variants A3 and C are considered rare. Interestingly, variant E was found exclusively in the Italian Piemontese breed, while variant F was found with a very low frequency (0.006) in the Emilia Romagna region (Northern Italy).

As shown in Fig. 6, variants A3, D, E, H1, H2 and I can be grouped together with variant A2 because they also have a proline at position 67 (Sebastiani et al., 2020, Noumin et al., 2022). Similarly, variants B, C, F and G, which have a histidine at position 67, can be considered as part of group A1.

This division into groups based on proline or histidine position provides a clearer view of the relationships between genetic variants in cattle and could have significant implications for research and population management.

Beta-casein variants	Change in amino acid sequence													
	18	25	35	36	37	67	72	88	93	106	117	122	137	138
A2	Ser-P	Arg	Ser-P	Glu	Glu	Pro	Glu	Leu	Gln	His	Gln	Ser	Leu	Pro
A1						His								
A3										Gln				
B						His						Arg		
C			Ser		Lys	His								
D	Lys													
E				Lys										
F						His								Leu
G						His						Leu		
H1		Cys						Ile						
H2							Glu		Leu					Glu
I									Leu					

Fig. 6. Changes in the amino acid sequence of beta casein variants (Kaminski et al., 2007)

In ovine beta casein, only the A and C genetic variants have been characterized at the protein level, showing a p.Glu2→Gln exchange. The genetic control of the B variant has not been assessed. Additionally, a transition from A to G in position 12029 along the exon 7, resulting in a p.Met183 → Val exchange, was identified at the DNA level and named variant G (Ceriotti et al., 2004). This variant was later described in a study by Corral et al. (2010) it was observed that the GG genotype was associated with increased milk production, while the AA genotype correlated with higher fat and protein percentages. Chessa et al. (2010) later identified two new patterns, X and Y. Upon sequencing, they found that the Y pattern resulted from a silent mutation (G→A transition) in the triplet coding for Gln at position 192, and the X variant was characterized by a C→A transversion leading to the amino acid exchange Leu196→Ile.

β-CN is a phosphoprotein of 207 amino acid residues in goats, which is shorter than the corresponding bovine protein (209 amino acids) because of the dipeptide deletion Pro179-Tyr180 (Richardson & Creamer, 1974). The gene consists of 9 exons ranging in size from 24 bp (exon 5) to 492 bp (exon 7). Exons 2 and 8 contain the translation start and stop signals, the former encoding the first two amino acids of the mature protein and the latter the last codon (Val). Exon 7, which is the longest, encodes approximately 82% of the mature protein and essentially corresponds to the hydrophobic part of the protein itself (Ramunno et al., 2007).

In goats, twelve allelic variants (A, A1, B, C, C1, C2, D, E, F, F1, G, 0, and 0') have been identified so far, responsible for the synthesis of eight protein variants. Variants *CSN2**A, *CSN2**B, *CSN2**C,

*CSN2**D, and *CSN2**E are considered "strong" variants associated with high β-CN content in milk (5 g/L per allele) (Roberts et al., 1992; Mahè and Grosclaude, 1993; Neveu et al., 2002; Galliano et al., 2004; Caroli et al., 2006). The variants *CSN2**C2 and *CSN2**F1 are associated with intermediate β-CN content (3.3 and 2.7 g/L per allele, respectively). The lower content of β-CN C2 and β-CN F1 variants in the milk suggests that C2 and F1 coding alleles, likely derived from C1 and A1 alleles, respectively, might be due to the presence of a C/T transition within the common non-coding ninth exon (Nicolai et al., 2021). On the other hand, alleles *CSN2* 0 and 01 are the so-called null alleles, as they are associated with negligible amounts of this protein in milk (Ramunno et al., 1995; Persuy et al., 1999).

In this case, as described for *CSN1S1*, allelic variants are characterized by individual SNPs (alleles A1, C, C1, C2, E, F, F1, G, and 0') or single deletions (allele 0). The other variants have been characterized only from a protein perspective (variants B and D) (Cosenza et al., 2005; Mahè and Grosclaude, 1993; Chessa et al., 2008; Nicolai et al., 2021; Galliano et al., 2004; Caroli et al., 2006; Chianese et al., 2007; Moatsou et al., 2007; Martin et al., 2013; Rahmatalla et al., 2021; Persuy et al., 1999; Rando et al., 1996; Cosenza et al., 2007; Cosenza et al., 2016). *CSN2**A is considered the ancestral allele, as it shares the amino acid alanine at position 177 with sheep, cattle, and buffalo (Chessa et al., 2005).

The study on beta casein variants in buffalo, initiated by Aschaffenburg and colleagues in 1968, marked a significant discovery. During a comparison of buffalo and cow milk, they identified two bands corresponding to β-CN. This led to the hypothesis of polymorphism in buffalo milk proteins.

Cosenza et al. in 2009 uncovered a single nucleotide polymorphism (SNP) at nucleotide 274 of exon 9, near the polyadenylation site, resulting in a silent allele named *CSN2* A1. They identified 13 mutations in exons, with three being non-conservative. Furthermore, they compared the promoter sequence with that of buffalo reared in India, revealing 25 mutations, with four potentially creating extra sites of regulation for *CSN2* gene expression in the Italian Mediterranean Buffalo.

Ramesha et al., (2016) analyzing Indian buffalo breeds observed only the A2 allele in the population. Pineda et al. (2019), with an extensive study of native Philippine swamp buffalo, found 4 polymorphic sites in CSN2, confirming the homozygosity already found by Ramesha et al. (2016) for the A2 variant in this species. De Oliveira et al in 2021 examined 657 buffaloes of four recognized breeds in Brazil, all confirmed to be A2A2. Additionally, a non-synonymous C>G SNP at position 78 of exon 7 was detected, changing lysine to asparagine, consistent with previous findings by Cosenza et al. (2009).

In 2023, Khan et al conducted a study on “Azi-Kheli” buffalo breed, finding all screened buffaloes to be homozygous for A2A2. Additionally, three new SNPs were identified (g.20545A > G, g.20570G

> A, and g.20693C > A) at positions 124, 160, and 760 in the exon 7, with the SNP g.760 C > A favourably associated with production traits. This SNP causes the amino acid change of threonine with Valine in the final protein and is associated with higher milk protein percentage, suggesting that this SNP could be used as a potential marker for milk production and quality improvement in buffalo.

CSN1S2

The *CSN1S2* gene is the last gene encoding calcium-sensitive caseins in order of position on the chromosome, giving rise to the protein αs2-CN, which is known to be highly phosphorylated. αs2-CN represents up to 10% of the casein fraction of bovine milk, in goats it reaches about 19% and in sheep it represents 22.8% of the total caseins (Selvaggi et al., 2014b). In buffalo milk, αs2-CN stands out as the third most abundant casein fraction, with a concentration of 4.99 g/L (Cosenza et al., 2021).

In cattle, 4 protein variants of αs2-casein are known (A, B, C and D). As reported by Selvaggi et al. (2014b), protein variant A was initially chemically determined and is the most frequent in all breeds investigated so far, being nearly fixed in most western breeds. Variants B and C are specific to zebu and yak, respectively. Variant D differs from αS2-CN A due to the deletion of 9 amino acid residues from positions 51 to 59 (Farrell et al., 2004). Variant C differs from variant A at positions 33, 47, and 130. The specific mutation sites leading to αS2-CN B have not been identified (Farrell et al., 2004). Additionally, Gallinat et al. (2013) identified a variant that they named *CSN1S2**E in a single animal. This variant is characterized by a nucleotide substitution c.64G > A, resulting in the amino acid exchange p.Val22→Ile in the final protein.

In sheep Boisnard, et al. (1991) reported the existence of two ovine genetic variants that, at the amino acid level, involve the replacement of amino acids, p.Asn49→Asp and p.Lys200→Asn respectively, this variants were lately observed also by Chessa et al. (2010) who named them as E and F variants. To date, ovine αs2-CN has seven variants named from A to G, with A considered the ancestral form. Chianese et al. (1993) identified three proteins due to their distinct electrophoretic behavior.

Chessa et al. (2003) recognized two phenotypes, named variant A and variant B, later characterized by Picariello et al. (2009) with two amino acid substitutions, p.Ile105 → Val and p.Asp75 → Tyr, as reported by Selvaggi et al. (2014b). Subsequently, Giambra and colleagues (2012) identified two additional variants, C and D, which were later characterized. *CSN1S2**C is characterized by two non-synonymous single nucleotide polymorphisms (SNPs) in exon 7 (c.178A>G, c.187G>T), leading to amino acid substitutions p.Val45→Ile and p.Ala48→Ser. *CSN1S2**D is caused by the SNP c.183G>C, resulting in an amino acid replacement at position 46 (p.Arg46→Ser).

In the same study, Giambra and colleagues reported a novel variant (variant G) characterized by c.527G>A within exon 15, resulting in the amino acid substitution p.Arg161→His. This particular

variant could not be identified using isoelectric focusing, and hence, it had not been reported until then. Furthermore, within the same study,

For this gene, 14 alleles responsible for as many protein variants have been identified in goat (Rhamatalla et al., 2022). Recently, Rahmatalla et al. (2021) identified four new variants, proposed as *CSN1S2*∗H, *CSN1S2*∗I, *CSN1S2*∗J, and *CSN1S2*∗K, following the existing alphabetical order for this protein, although the corresponding quantity of αs2-CN has not yet been investigated.

The alleles associated with a high content of αs2 are *CSN1S2**A, *CSN1S2**B, *CSN1S2**C, *CSN1S2**E, *CSN1S2**F (Boulanger et al., 1984; Grosclaude et al., 1987; Bouniol et al., 1994; Lagonigro et al., 2001; Ramunno et al., 2001a, Cunsolo et al., 2006), while the *CSN1S2**D allele contributes to an intermediate content (Ramunno et al., 2001a). The *CSN1S2**0 allele is considered a "null" allele, associated with an apparent absence of αs2-CN (Ramunno et al., 2001b). Furthermore, Erhardt et al. (2002) reported the existence of a variant named G, associated with a normal content of αs2-CN but observable exclusively through isoelectric focusing. Finally, Cunsolo et al. (2006) noted the presence of two truncated protein variants identified as truncated sub-variant A and truncated sub-variant E.

*CSN1S2**A is considered the ancestral form in goat, following the evolutionary path from sheep and cattle (Sacchi et al., 2005). In this case as well, allelic variants are characterized by nucleotide substitutions (B, C, E, F, H, I, J, and K) or deletions (variants D and 0) (Boulanger et al., 1984; Bouniol et al., 1994; Ramunno et al., 2001a; Chianese et al., 1998; Veltri et al., 2000; Lagonigro et al., 2001; Erhardt et al., 2002; Rahmatalla et al., 2021; Rahmatalla et al., 2022; Ramunno et al., 2001b). Little is known at the nucleotide level about variant G, and even less about the two truncated proteins A and E, although it is hypothesized that they originate from a differential splicing process of pre-mRNA during the translation process of *CSN1S2**A and *CSN1S2**E variants (Cunsolo et al., 2006).

In the case of the buffalo, the αs2-CN had never been characterized in detail until 2021, and in fact the knowledge about the allelic variations were limited to variants A and B. Cosenza et al., (2011) identified the allele named B distinguished by the g.773G>C mutation. This mutation causes the inactivation of the splice donor site in intron 7, consequently triggering the skipping of the exon. Further insights through cDNA sequencing revealed three exon mutations, with the first, c.459C>T, being silent, while the other two, c.484A>T and c.568A>G, caused a non-conservative change of amino-acids p.Ile162→Phe and p.Thr200→Ala, respectively. Subsequently, in 2021, Cosenza and colleagues confirmed the existence of these three SNPs, suggesting the presence of at least eight

allelic variants of the *CSN1S2* gene (Fig. 7), responsible for seven different possible translations of buffalo αs2-CN.

CSN1S2 alleles	Exon, nucleotide, and amino acid position																Breed
	Exon 2		Exon 7		Exon 11		Exon 11		Exon 13		Exon 14		Exon 16		Exon 16		
	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa	
	3165	5	7539	58-66	9220	127	9221	128	11072	153	12803	162	14067	190	14141	214	
A[1]	T	I	G	EVIRNANEE	T	V	A	K	C	T	T	F	G	A	C	N	Murrah
C[2]	T	I	G	EVIRNANEE	T	V	A	K	C	T	A	I	G	A	C	N	Mediterranean
D[3]	T	I	G	EVIRNANEE	T	V	A	K	C	T	A	I	A	T	C	N	Mediterranean/ Egyptian/ Murrah
E[4]	C	I	G	EVIRNANEE	T	V	A	K	C	T	A	I	A	T	G	K	Murrah
F[5]	T	I	G	EVIRNANEE	A	V	G	E	C	T	A	I	A	T	C	N	Murrah
B[6]	T	I	C	—	T	V^{117}	A	K^{118}	T	T^{144}	T	F^{153}	G	A^{181}	C	N^{205}	Mediterranean
B1[7]	T	I	C	—	T	V^{117}	A	K^{118}	C	T^{144}	T	F^{153}	G	A^{181}	C	N^{205}	Carabao
B2[8]	T	I	C	—	T	V^{117}	A	K^{118}	C	T^{144}	A	I^{153}	A	T^{181}	C	N^{205}	Mediterranean

Fig. 7. Genetic variants of buffalo αs2-CN casein encoding gene (Cosenza et al., 2021)

The allele named *CSN1S2* A, identified in the study of Cosenza and al (2021), is considered the ancestral form of αs2-CN, in accordance with the nucleotide and amino acid sequences found in bovines and goats. Cosenza and colleagues in 2021 propose two potential evolutionary pathways characterized by mutational events, amino acid substitutions, and deletions, starting from the ancestral form *CSN1S2* A (the allele that Cosenza et al 2021 called D was called the A allele in the earlier studies, so the A allele that Cosenza et al 2021 refer to is different from that in the previous studies). The first evolutionary map gives rise to four distinct alleles (*CSN1S2* C, D, E, F), each characterized by specific amino acid modifications. The second map, generated by the point mutation g.7539G>C, leads to the inactivation of the splice donor site in intron 7, giving rise to alleles *CSN1S2* B1 and B. These alleles are distinguished by the complete skipping of exon 7, with further variation in the single nucleotide polymorphism g.11072C>T in exon 13 (Fig. 8).

Within the same study, a significant analysis was conducted on the relationships between *CSN1S2* polymorphisms and the fatty acid profile.

A particular focus was placed on the SNP g.14067A>G, showing a significant impact ($P < 0.05$) on the palmitic acid content in buffalo milk. Buffaloes with GG and heterozygous genotypes exhibited a reduction of 34.13% and 34.71%, respectively, compared to the AA genotype (35.23%). These results provide an intriguing perspective on the genotype-phenotype correlation in the context of the lipid composition of buffalo milk.

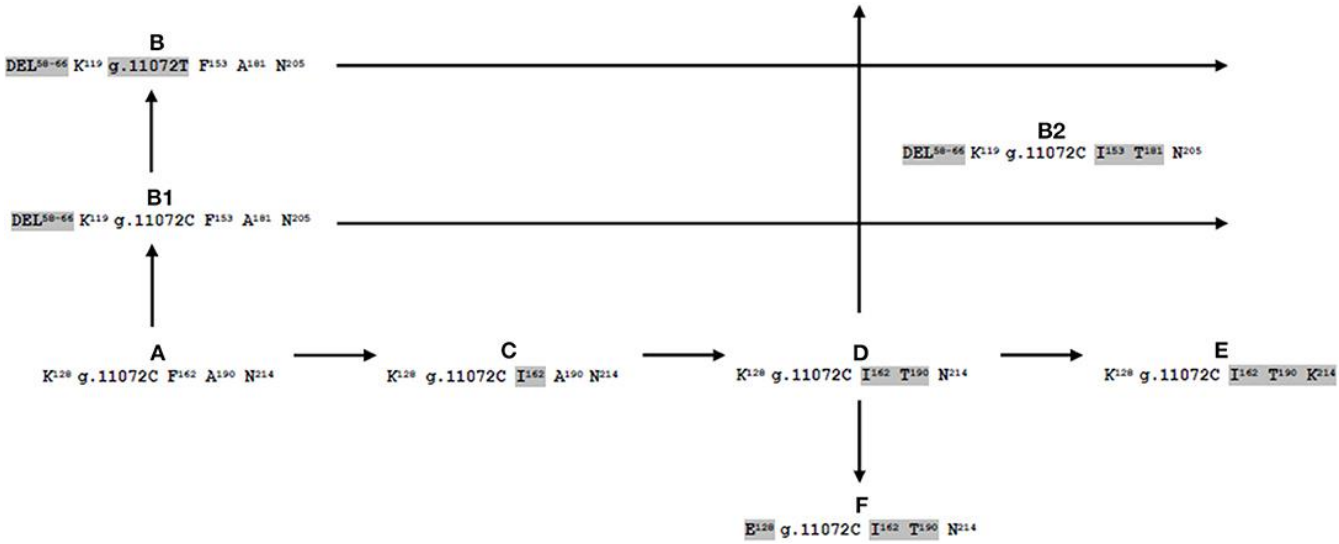

Fig. 8. Possible evolutionary pathway of buffalo αs2-casein-encoding gene (*CSN1S2*) (Cosenza et al., 2021).

DGAT1

Acyl-CoA: diacylglycerol acyltransferase (DGAT) is a microsomal enzyme that plays a central role in glycolipid metabolism at the cellular level, catalyzing the final step of triacylglycerol synthesis using diacylglycerol (DAG) as a substrate and acyl carbon fatty acid (Bell and Coleman, 1980; Lehner and Kuksis, 1996). DGAT1 belongs to a large family of membrane-bound O-acyltransferases (MBOATs) and membrane-associated enzymes that catalyze O-acylation reactions by transferring fatty acyl groups to hydroxyl or thiol groups of lipid and protein receptors. Its members are involved in lipid metabolism, signal transduction, and hormone processing; a common feature of the MBOAT family is a long hydrophobic region containing asparagine and histidine residues in the putative active site (corresponding to amino acids 378 and 415 in human DGAT1; 389 and 426 in mouse DGAT1). DGAT is thought to play a major role in intestinal fat absorption (Brindley, 1991), regulation of blood plasma triglyceride concentration (Bell and Coleman, 1980; Haagsman and Van Golde, 1981), fat accumulation in adipocytes (Coleman and Bell, 1976), and energy metabolism in muscle cells (Swanton and Saggerson, 1997). DGAT also plays an important role in oil synthesis in seeds (Lassner, 1997), triglyceride synthesis in eggs (Homa et al., 1986), and milk (Bell and Coleman, 1980).

To date, the *DGAT1* gene has been sequenced in several species of zootechnical interest: pig (Nonneman and Rohrer, 2002, GenBank accession number AY116586), cattle (Winter et al., 2002, GenBank accession number AJ318490; Grisart et al., 2002 GenBank accession number AY065621), buffalo (Yuan et al., 2007; GenBank accession number AY999090.1), goat (GenBank accession number LT221856.1), and sheep (Scatà et al., 2009; GenBank accession number EU178818.1).

The bovine *DGAT1* gene spans a DNA tract of approximately 8303 bp, including 1742 bp of exonic regions and 6561 bp of intronic regions. The main feature of the bovine *DGAT1* gene is its extremely fractionated architecture. This gene consists of 17 exons, ranging in size from 39 bp (exon 10) to 432 bp (exon 17), and 16 introns. This structure is similar to that of the pig and human *DGAT* genes, both in terms of the size of the introns and the organization of the 17 exons (Cases et al., 1998; Smith et al., 2000). *DGAT1* is located on bovine chromosome 14 in the region of a QTL associated with improved milk fat production. The area containing this QTL lies between the BULGE13-BULGE9 microsatellite markers, and the data obtained by Grisart et al. (2001) suggest that *DGAT1* is the candidate gene associated with this QTL. In the buffalo species, the *DGAT1* gene (10.2 kp) is located on chromosome 15 and is characterized by 17 exons (NC_037559), similar to that in cattle, while in the goat species, it is located on chromosome 14 (https://www.ncbi.nlm.nih.gov/gene/100861225). Similarly, in sheep, the *DGAT1* gene consists of 17 exons spanning 8.67 kb on chromosome 9 and encodes a protein of approximately 489 amino acids.

Bovine

In cattle, at the *DGAT1* locus, an interesting polymorphism has been detected at the protein level, where a non-conservative substitution of lysine (K) by alanine (A) (K232A) is observed at position 232. This polymorphism appears to have a significant effect on the fat content in the milk and other milk characteristics in Jersey, Ayrshire, Holstein New Zealand, Dutch, German, Israeli, and Polish breeds (Grisart et al., 2002; Spelman et al., 2002; Thaller et al., 2003; Weller et al., 2003). The K allele exhibits a higher VMAX response than the A allele and is also associated with higher milk fat content (+0.35%) and increased milk production (+10 kg) (Grisart et al., 2004; Spelman et al., 2002; Winter et al., 2002).

From the analysis of the sequences present in the database (GenBank accession number AW446985), the existence of alternative splicing has been demonstrated. This involves the appearance of two transcripts of different lengths. The longer one utilizes the end of the eighth intron, 6 bp upstream of the polymorphic site K232A, resulting in the intronification of the eighth exon. By amplifying the region containing the alternative splicing site, it was possible to verify that the shortest variant was associated with the K allele of the K232A locus (Grisart et al., 2004). Subsequent studies have shown the existence of other polymorphisms besides K232A: two NSPs (A G; C T) at intron level 12 [N984+8(AG) and N984+26(CT)] and a SNP (C T) in the 3'UTR region (Nt-1501) (Grisart et al., 2001; Secchiari et al., 2009).

In addition, a VNTR (variable number of tandem repeats) (5'-AGGCCCCGCCCTCCCCGG-3') located in the region of the bovine *DGAT1* gene promoter, whose sequence contains the binding site (CCCCCGCC) for the specificity protein 1 gene regulator (Sp1), was characterized. The authors hypothesized that alleles with a greater number of repetitions of the 18 bp sequence may increase the expression of *DGAT1* in the mammary gland, resulting in higher fat production in milk (Kühn et al., 2004).

Sheep

In sheep, the *DGAT1* gene (gene ID: 100126245) is located on chromosome 9 and consists of 17 exons. The gene has been almost completely sequenced by Scatà et al. (2009) (GenBank accession number EU178818) in three Italian breeds: Altamurana, Gentile di Puglia, and Sarda. The gene spans 8670 bp and encodes a 489 aa protein with 97% nucleotide identity to the bovine homologue. At the protein level, the sequence of the two species differs by 5 amino acid changes: p.Arg26Gly in exon 2; p.Leu387Ile in exon 14; p.Gly407Ala in exon 15; p.Arg425Gly in exon 16, and p.Thr486Ala in exon 17. The variant p.Lys232Ala, which characterizes the bovine at the level of exon 8 and is

responsible for the variability of the fat content in milk, does not seem to characterize the ovine *DGAT1*, which would be monomorphic due to the presence of the lysine residue.

Regarding the study of polymorphisms, the same authors report the identification of 5 localized SNPs: in the promoter region (5'UTR, EU178818: g.127C>A, which could affect a binding site of the gene regulator Sp1), in intron 1 (g.1655C>T), intron 2 (g.5553C>T), intron 10 (g.7492C>T), and a non-synonymous mutation (p.Ala487) in exon 17 (g.8539C>T).

A subsequent study by Dervishi et al. (2015) in Spanish sheep of the Assaf breed showed three additional polymorphisms: g.358C>A in Sone 1, responsible for the aspartic acid change to glutamic acid at position 53, g.7457C>A in intron 10, and g.8522C>T in Sone 17, responsible for the change aa p.Arg482Cys. To date, several association studies have been carried out between some of these markers and the qualitative-quantitative variability of the lipid composition of sheep's milk (Scatà et al., 2009; Dervishi et al., 2015).

Goat

The *DGAT1* gene in the caprine species has been little studied, and there have been few investigations to identify polymorphisms useful for exploring associations with traits of interest. The first example of a polymorphism was reported by Angiolillo et al. (2007), who described a T-C transition in intron 16. Later, An et al. (2013) identified several other polymorphisms at this locus: g.407-408insC (intron 14), g.6852C T, and g.6798C T (in exon 7) in the Xinong Saanen (SN) and Guanzhong (GZ) goat breeds and reported a statistically significant association with milk production and fat percentage. More recently, Martin et al. (2017) described 29 polymorphic sites at the *DGAT1* locus in goats reared in France. Of these, two are responsible for amino acid changes: R251L and R396W. Both mutations were associated with significant variability in milk fat content.

Buffalo

Buffalo milk is characterized by a higher milk fat content than that of other ruminants of zootechnical interest. Given the importance of this trait, several studies have been conducted to identify polymorphisms at the *DGAT1* locus also in this species and possible associations with the observed phenotypic-quantitative variability. Gu et al. (2017) identified a 69-nucleotide silent SNP in Sone 13 and 10 polymorphic intronic sites. Li et al. (2018) describe a SNP (g.9046T > C) located in Sone 17 and responsible for an arginine-histidine amino acid change, which is significantly associated with fat percentage: genotype TT is associated with a significantly higher fat percentage than genotype CC. Previously, De Freitas et al. (2016) always described a SNP (g.11785 T> C) at Sone 17 responsible for the change aa p.Ala484>Val, significantly associated with the percentage of fat and

milk proteins in Murrah buffalo ($P<0.05$). Contrary to the results reported by De Freitas et al. (2016), Meng et al. (2013) show no correlation between polymorphisms localized in exon 17 and milk production and composition in buffaloes of the Swamp genus. Similarly to cattle, Cardoso et al. (2015) observed a significant association between the variability of the VNTR in the promoter region of the buffalo *DGAT1* gene and the percentage of fat in milk.

MBL

The *MBL* gene, encoding for mannose-binding lectin (MBL), is an acute-phase protein that binds to mannose and N-acetyl-D-glucosamine sugars present on a wide range of pathogens, including gram-positive and gram-negative bacteria, yeasts, parasites, and viruses (Turner 1998; Takahashi et al., 2005). This binding on microbial surface may cause the direct absorption of the microbial agent from the phagocytes or the activation of a pro-serin proteases complex formed by MASP 1 e MASP 2 leading to the splitting of C4 and C2 of the complement system. (Thiel et al. 1997, Nepomuceno et al. 1997. Turner 1998)

The antimicrobial functions of MBL depend on the correct oligomeric assembly and clustered orientation of carbohydrate recognition domains. These domains recognize and bind to molecular patterns associated with pathogens repeated on microbial surfaces (van de Wetering et al. 2004).

Most mammalian species possess *MBL1* and *MBL2* genes, encoding for MBL-A and MBL-C proteins, respectively. Studies have highlighted those mutations in both genes contribute to variation in susceptibility to various infections (Capparelli et al., 2008, Lilli et al., 2005, Shi et al., 2004). Although *MBL1* has been identified as a pseudogene in humans and chimpanzees, MBL-A has been isolated in pigs and cattle (Lillie et al., 2005, Lillie et al., 2006, Liu et al., 2011, Wang et al., 2011). In the study by Liu and colleagues (2011) on cattle, similarities were detected between the promoter region of the porcine *MBL1* gene and that of cattle, highlighting the connection between bovine and porcine *MBL1* genes. The presence of *MBL1* was studied in cows, both healthy and affected by mastitis, with higher mRNA levels in the liver tissues of mastitic cows. Wang et al. (2011) also suggested that a single nucleotide polymorphism (SNP) on exon 2 of the *MBL1* gene was associated with the somatic cell score. The *MBL2* gene in ruminant species such as cattle, buffalo, sheep, and goats exhibit common structural features. It varies in size from 5105 bp (buffalo) to 6755 bp (goat), with 4 introns and 5 exons. Remarkably, exon 1 is non-coding, and its chromosomal position varies among species: in cattle and goat, it is on chromosome 26; in buffalo, on chromosome 23; and in sheep, on chromosome 22. (https://www.ncbi.nlm.nih.gov/gene/281297; https://www.ncbi.nlm.nih.gov/gene/?term=mbl2+buffaslo; https://www.ncbi.nlm.nih.gov/gene/102181921; https://www.ncbi.nlm.nih.gov/gene/101116362).

As studies on the genetic variability of the *MBL2* gene in ruminants are rather limited, the structure and the most important polymorphisms of the human *MBL2* gene are shown in Fig. 9.

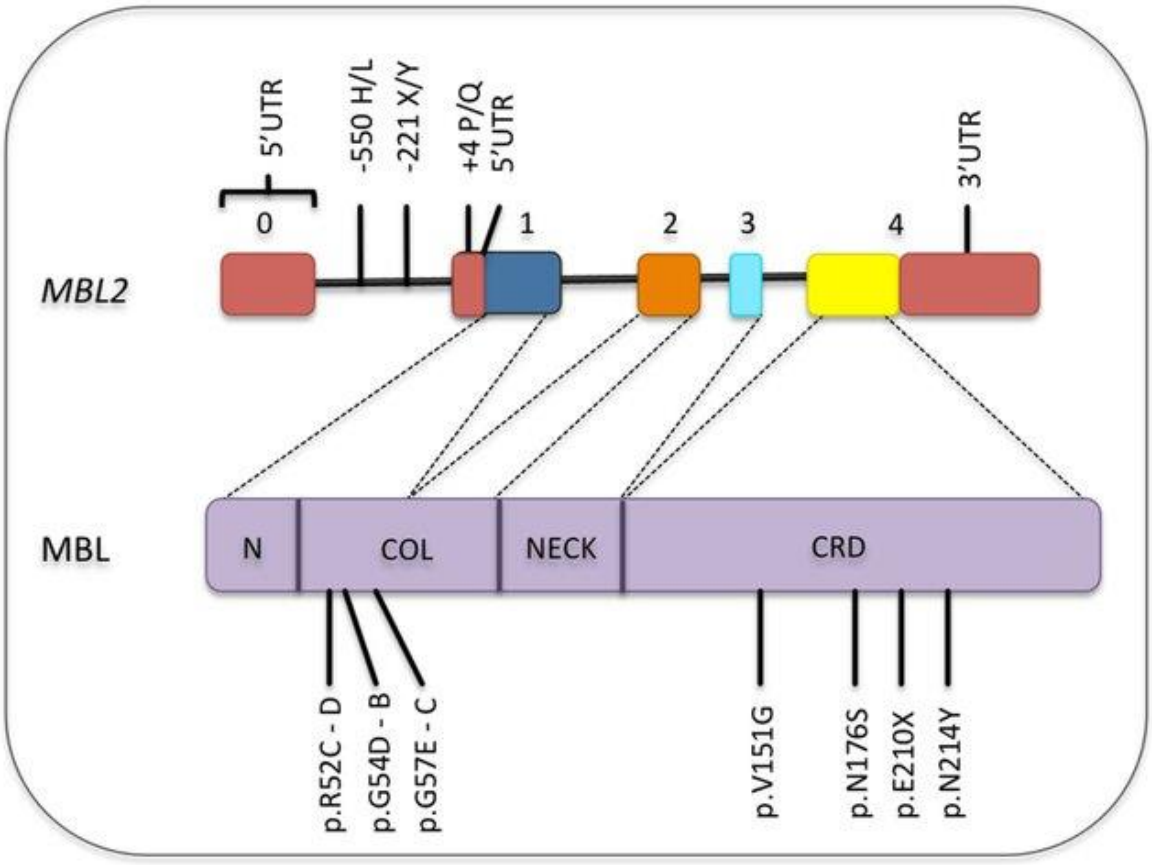

Fig. 9. Common polymorphisms in the human *MBL2* gene and its corresponding locations in the MBL protein. Only the functional polymorphisms in the promoter and non-synonymous mutations are shown. Exons are numbered. Exons, introns, and protein domains are not in scale. N, N-terminal region; COL, collagen-like region; CRD, carbohydrate-recognition domain. (Beltrame et al., 2015).

Nevertheless, two studies conducted by Zhao et al. (2011) and Cosenza et al. (2012) on sheep provided important information for this species.

Zhao et al. (2011) examined the exon 1 region of the *MBL*2 gene in Hu sheep, identifying four distinct patterns. Pattern D remained unchanged compared to reference sequences, while pattern C had two SNPs (c.115C>T and c.53G>A), with the latter causing a variation in the aminoacidic sequence Val18→Met. Pattern B exhibited three SNPs at different positions (c.37, c.51, and c.93), of which two (c.52 and c.93) were non-synonymous, leading to amino acid substitutions p.Glu17→Val and p.Glu31→Gly, respectively. In contrast, pattern A was significantly different, characterized by a deletion of 3bps between the 38th and 42nd bases, another deletion of 6bps between the 84th and 91st bases, and a total of 42 SNPs.

The presence of these patterns influences the expressed protein quantity, with patterns A and B associated with lower levels compared to patterns C and D. It is assumed that this variation in MBL-C levels is associated with a higher susceptibility of individuals carrying patterns A and B to infectious diseases (Capparelli et al., 2008; Lilli et al., 2005; Shi et al., 2004).

Cosenza et al. (2012) examined the promoter region and exon 1 of the *MBL*2 gene in Sarda sheep. This analysis revealed the presence of 13 polymorphic sites, with seven located in the first exon, five of which caused amino acid alterations. In the same study, Nicastrese goats were also investigated, extending up to nucleotide 417 of the fourth exon, revealing 12 polymorphic sites, with six distributed in introns 1, 2, and 3, five in the promoter region, and one in the second exon.

It's important to note that this study differed from the previous one conducted by Zhao et al. (2011). Cosenza et al. (2012) did not associate these variants with the corresponding protein quantity (MBL-C). Consequently, it's challenging to understand if these SNPs could be responsible for increased susceptibility to infectious diseases, unlike the conclusions drawn in previous works (Capparelli et al., 2008; Lilli et al., 2005; Shi et al., 2004).

In contrast to goat and sheep species, studies on cattle and buffalo have revealed a significant correlation between certain SNPs along the *MBL*2 gene and milk quality parameters, as well as the Somatic Cell Score (SCS), an indicator of mastitis. In particular, Wang et al. (2012) identified 4 SNPs along exon 1 of the MBL2 gene in Holstein and Luxi yellow cows. Although two of these SNPs (g.201 G>A and g.234 C>A) were not directly associated with the MBL-C protein quantity, they were significantly correlated with the SCS.

These results were subsequently confirmed by Zhao et al. (2012), who identified the same SNPs found by Wang. In this case, the association was highlighted not only with the SCS (g.1164 G>A, g.1197 C>A) but also with the fat and protein content of milk (g.1164 G>A). The repetition of these results in two distinct studies suggests a potential role of the *MBL*2 gene in mastitis susceptibility in cattle, emphasizing the importance of further investigations to fully understand the involvement of these genetic variants in processes related to mammary health in cattle.

Concerning river buffaloes, two relevant studies have primarily been conducted, the first from Capparelli et al., (2008) and the second from Shregojry et al. (2023).

Capparelli et al. (2008) identified buffalo haplotypes according to human nomenclature which is based on three sites of variability: at position -550 bp from exon 1 corresponding to H and L variant, respectively, at position -221 bp from exon 1 corresponding to X and Y variant, respectively and four possible alleles in exon 1 (A, B, C or D) thus giving rise to 16 possible haplotypes (HXA; HXB; HXC; HXD; HYA; HYB; HYC; HYD; LXA; LXB; LXC; LXD; LYA; LYB; LYC; LYD). Capparelli et al. (2008), showed that individuals homozygous for the HYA haplotype were resistant to infections by *Brucella abortus*. In contrast, individuals homozygous for the LYD haplotype were susceptible to the pathogen, providing important information on the correlation between genetic variants of the *MBL*2 gene and resistance/susceptibility to specific pathogens.

A recent study conducted by Shregojry et al. (2023) on Murrah-type buffaloes identified four SNPs (1262G>A, 3382A>T, 4387C>T, and 4511C>T) of the *MBL2* gene (NCBI GenBank AC_000183.1) associated with an increased risk of clinical mastitis. Variants A, A, C and C at positions 1262, 3382, 4387 and 4511 along the gene were more frequently associated with mastitis-affected animals than unaffected ones, identifying these alleles as 'at risk'. Genotype analysis confirmed these findings, revealing that Murrah buffaloes with AG or GG genotype at position 1262G>A had significantly lower clinical mastitis compared to the AA genotype ($P<0.01$). Similarly, genotypes AT or TT at position 3382A>T, TC or TT at position 4387C>T, and TT or CT at position 4511C>T were associated with a significant reduction in the risk of clinical mastitis compared to the respective reference genotypes AA, CC, and CC ($P<0.01$).

In summary, the limited availability of studies exploring the correlation between *MBL2* gene SNPs and mastitis underscores the need for a deeper understanding of the genetic aspects of mastitis in ruminant species. The current research landscape indicates the presence of a network of associations between genetic variants of *MBL2* and susceptibility to infections, as well as milk quality parameters such as the Somatic Cell Score (SCS). These clues suggest a possible connection between genetic variants of this gene and the incidence of mastitis in the examined populations.

It is important to undertake further studies to confirm and deepen what has been discovered so far, providing a solid foundation for fully understanding the role of the *MBL2* gene in susceptibility to mammary diseases in ruminants. The prospect of increased genetic knowledge could pave the way for more targeted selection strategies and preventive interventions aimed at improving resistance to mastitis and, consequently, the overall health of cattle herds and flocks of sheep and goats.

References

1. An, X. P.; Hou, J. X.; Zhao, H. B.; Bai, L.; Peng, J. Y.; Zhu, C. M.; Yan, Q. M.; Song, Y. X.; Wang, J. G.; Cao, B. Y. Polymorphism Identification in Goat DGAT1 and STAT5A Genes and Association with Milk Production Traits. Czech J. Anim. Sci. 2013, 58 (7), 321–327. https://doi.org/10.17221/6862-CJAS.
2. Angiolillo, A.; Amills, M.; Urrutia, B.; Doménech, A.; Sastre, Y.; Badaoui, B.; Jordana, J. Identification of a Single Nucleotide Polymorphism at Intron 16 of the Caprine Acyl-Coenzyme A: Diacylglycerol Acyltransferase 1 (*DGAT1*) Gene. *Journal of Dairy Research* **2007**, *74* (1), 47–51. https://doi.org/10.1017/S0022029906002196.
3. Anusha Siddiqui, S.; Mahmood Salman, S. H.; Ali Redha, A.; Zannou, O.; Chabi, I. B.; Oussou, K. F.; Bhowmik, S.; Nirmal, N. P.; Maqsood, S. Physicochemical and Nutritional Properties of Different Non-Bovine Milk and Dairy Products: A Review. *International Dairy Journal* **2024**, *148*, 105790. https://doi.org/10.1016/j.idairyj.2023.105790.
4. Argyriadou, A.; Michailidou, S.; Vouraki, S.; Tsartsianidou, V.; Triantafyllidis, A.; Gelasakis, A.; Banos, G.; Arsenos, G. A Genome-Wide Association Study Reveals Novel SNP Markers Associated with Resilience Traits in Two Mediterranean Dairy Sheep Breeds. *Front. Genet.* **2023**, *14*, 1294573. https://doi.org/10.3389/fgene.2023.1294573.
5. Aschaffenburg, R.; Sen, A.; Thompson, M. P. The Caseins of Buffalo Milk. *Comparative Biochemistry and Physiology* **1968**, *27* (2), 621–623. https://doi.org/10.1016/0010-406X(68)90262-4.
6. Balteanu VA, Pop FD, Vlaic A, Rusu AR (2008) Multiple mutations events are characterizing alpha s1 casein Brv and beta casein Crv alleles discovered in Romanian buffalo breed. *Bulletin of University Of Agricultural Sciences And Veterinary Medicine Cluj-Napoca. Animal Science and Biotechnologies* **65**, 478.
7. Bell, R. M.; Coleman, R. A. Enzymes of Glycerolipid Synthesis in Eukaryotes. *Annu. Rev. Biochem.* **1980**, *49* (1), 459–487. https://doi.org/10.1146/annurev.bi.49.070180.002331.
8. Beltrame, M. H.; Catarino, S. J.; Goeldner, I.; Boldt, A. B. W.; De Messias-Reason, I. J. The Lectin Pathway of Complement and Rheumatic Heart Disease. Front. Pediatr. 2015, 2. https://doi.org/10.3389/fped.2014.00148.
9. Boisnard, M.; Hue, D.; Bouniol, C.; Mercier, J.; Gaye, P. Multiple mRNA Species Code for Two Non-allelic Forms of Ovine As2-casein. *European Journal of Biochemistry* **1991**, *201* (3), 633–641. https://doi.org/10.1111/j.1432-1033.1991.tb16324.x.
10. Bonfatti, V.; Gervaso, M.; Rostellato, R.; Coletta, A.; Carnier, P. Protein Composition Affects Variation in Coagulation Properties of Buffalo Milk. *Journal of Dairy Science* **2013**, *96* (7), 4182–4190. https://doi.org/10.3168/jds.2012-6333.
11. Bonfatti, V.; Giantin, M.; Gervaso, M.; Coletta, A.; Dacasto, M.; Carnier, P. Effect of CSN1S1-CSN3 (αS1-κ-Casein) Composite Genotype on Milk Production Traits and Milk Coagulation Properties in Mediterranean Water Buffalo. *Journal of Dairy Science* **2012a**, *95* (6), 3435–3443. https://doi.org/10.3168/jds.2011-4901.
12. Bonfatti, V.; Giantin, M.; Gervaso, M.; Rostellato, R.; Coletta, A.; Dacasto, M.; Carnier, P. Short Communication: CSN1S1-CSN3 (αS1-κ-Casein) Composite Genotypes Affect Detailed Milk Protein Composition of Mediterranean Water Buffalo. *Journal of Dairy Science* **2012b**, *95* (11), 6801–6805. https://doi.org/10.3168/jds.2012-5601.
13. Boulanger, A.; Grosclaude, F.; Mahé, M.-F. Polymorphisme Des Caséines As1 et As2 de La Chèvre (Capra Hircus). *Genetics Selection Evolution* **1984**, *16* (2), 157. https://doi.org/10.1186/1297-9686-16-2-157.
14. Bouniol, C.; Brignon, G.; Mahé, M. F.; Printz, C. Biochemical and Genetic Analysis of Variant C of Caprine α_{s2} -casein (Capra Hircus). *Animal Genetics* **1994**, *25* (3), 173–177. https://doi.org/10.1111/j.1365-2052.1994.tb00106.x.
15. Brindley, D. N. (1991) in *Biochemistry of Lipids, Lipoproteins and Membranes* (Vance, D. E., and Vance, J. E., eds) pp. 171-203, Elsevier, Amsterdam
16. Capparelli, R.; Parlato, M.; Amoroso, M. G.; Roperto, S.; Marabelli, R.; Roperto, F.; Iannelli, D. Mannose-Binding Lectin Haplotypes Influence Brucella Abortus Infection in the Water Buffalo (Bubalus Bubalis). *Immunogenetics* **2008**, *60* (3–4), 157–165. https://doi.org/10.1007/s00251-008-0284-4.
17. Cardoso, D. F.; De Souza, G. F. P.; Aspilcueta-Borquis, R. R.; Araujo Neto, F. R.; De Camargo, G. M. F.; Hurtado-Lugo, N. A.; Scalez, D. C. B.; De Freitas, A. C.; Albuquerque, L. G.; Tonhati, H. Short Communication: Variable Number of Tandem Repeat Polymorphisms in DGAT1 Gene of Buffaloes (Bubalus Bubalis) Is Associated with Milk Constituents. *Journal of Dairy Science* **2015**, *98* (5), 3492–3495. https://doi.org/10.3168/jds.2014-8729.
18. Caroli, A. M.; Chessa, S.; Erhardt, G. J. Invited Review: Milk Protein Polymorphisms in Cattle: Effect on Animal Breeding and Human Nutrition. *Journal of Dairy Science* **2009**, *92* (11), 5335–5352. https://doi.org/10.3168/jds.2009-2461.
19. Caroli, A.; Chessa, S.; Chiatti, F.; Rignanese, D.; Meléndez, B.; Rizzi, R.; Ceriotti, G. Short Communication: Carora Cattle Show High Variability in As1-Casein. *Journal of Dairy Science* **2008**, *91* (1), 354–359. https://doi.org/10.3168/jds.2007-0420.
20. Caroli, A.; Chiatti, F.; Chessa, S.; Rignanese, D.; Bolla, P.; Pagnacco, G. Focusing on the Goat Casein Complex. *Journal of Dairy Science* **2006**, *89* (8), 3178–3187. https://doi.org/10.3168/jds.S0022-0302(06)72592-9.
21. Cases, S.; Smith, S. J.; Zheng, Y.-W.; Myers, H. M.; Lear, S. R.; Sande, E.; Novak, S.; Collins, C.; Welch, C. B.; Lusis, A. J.; Erickson, S. K.; Farese, R. V. Identification of a Gene Encoding an Acyl CoA:Diacylglycerol Acyltransferase, a Key Enzyme in Triacylglycerol Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (22), 13018–13023. https://doi.org/10.1073/pnas.95.22.13018.

22. Ceriotti, G.; Chessa, S.; Bolla, P.; Budelli, E.; Bianchi, L.; Duranti, E.; Caroli, A. Single Nucleotide Polymorphisms in the Ovine Casein Genes Detected by Polymerase Chain Reaction-Single Strand Conformation Polymorphism. *Journal of Dairy Science* **2004**, *87* (8), 2606–2613. https://doi.org/10.3168/jds.S0022-0302(04)73386-X.
23. Chessa S, Bolla P, Dario C, Pieragostini E, Caroli A (2003) Polimorfismi genetici lattoproteici nella razza ovina gentile di puglia: Monitoraggio mediante focalizzazione isoelectrica. Sci Tecn Latt Cas 54:1–8
24. Chessa, S.; Budelli, E.; Chiatti, F.; Cito, A. M.; Bolla, P.; Caroli, A. Short Communication: Predominance of β-Casein (CSN2) C Allele in Goat Breeds Reared in Italy. *Journal of Dairy Science* **2005**, *88* (5), 1878–1881. https://doi.org/10.3168/jds.S0022-0302(05)72863-0.
25. Chessa, S.; Rignanese, D.; Berbenni, M.; Ceriotti, G.; Martini, M.; Pagnacco, G.; Caroli, A. New Genetic Polymorphisms within Ovine β- and αS2-Caseins. *Small Ruminant Research* **2010**, *88* (2), 84–88. https://doi.org/10.1016/j.smallrumres.2009.12.014.
26. Chessa, S.; Rignanese, D.; Küpper, J.; Pagnacco, G.; Erhardt, G.; Caroli, A. Short Communication: The β-Casein (CSN2) Silent Allele C1 Is Highly Spread in Goat Breeds. *Journal of Dairy Science* **2008**, *91* (11), 4433–4436. https://doi.org/10.3168/jds.2008-1228.
27. Chianese L., Caira S., Garro G., Quarto M., Mauriello R., Addeo F. Occurrence of genetic polymorphism at goat apha-s1-CN locus. The challenge to sheep and goats milk sectors Alghero (Sassari) Italy 18-20 aprile **2007**. pp. 55-57. ISBN 978-9-290980-40-7
28. Chianese, L. (1997). The casein variants of ovine milk and the relationships between the αs1casein variants and milk composition, micellar size and cheese yield. Casein and Caseinates: Structures. In Interactions, Networks. Hannach Symposium, Scotland, UK.
29. Chianese, L., Garro, G., Ferranti, P., Speranza, L. M., Pilla, F., Bevilacqua, C., et al. (1998). "Occurence of new variant ofgoat αs1-and αs2-casein in abreed of Southern Italy," in Proceeding of 25th International Dairy Congress, 82nd annual sessions, Arahus, Denmark.
30. Chianese, L.; Garro, G.; Addeo, F.; Lopez-Galvez, G.; Ramos, M. Discovery of an Ovine α_{s2} -Casein Variant. *Journal of Dairy Research* **1993**, *60* (4), 485–493. https://doi.org/10.1017/S0022029900027849.
31. Chianese, L.; Quarto, M.; Pizzolongo, F.; Calabrese, M. G.; Caira, S.; Mauriello, R.; De Pascale, S.; Addeo, F. Occurrence of Genetic Polymorphism at the As1-Casein Locus in Mediterranean Water Buffalo Milk. *International Dairy Journal* **2009**, *19* (4), 181–189. https://doi.org/10.1016/j.idairyj.2008.09.008.
32. Coleman R, Bell RM. Triacylglycerol synthesis in isolated fat cells. Studies on the microsomal diacylglycerol acyltransferase activity using ethanol-dispersed diacylglycerols. J Biol Chem. 1976 Aug 10;251(15):4537-43. PMID: 947894.
33. Corral, J. M.; Padilla, J. A.; Izquierdo, M. Associations between Milk Protein Genetic Polymorphisms and Milk Production Traits in Merino Sheep Breed. Livestock Science 2010, 129 (1), 73–79. https://doi.org/10.1016/j.livsci.2010.01.007.
34. Cosenza G, Gallo D, Auzino B, Gaspa G, Pauciullo A. Complete CSN1S2 Characterization, Novel Allele Identification and Association With Milk Fatty Acid Composition in River Buffalo. Front Genet. **2021** Feb 4;11:622494. doi: 10.3389/fgene.2020.622494. PMID: 33613624; PMCID: PMC7890360.
35. Cosenza, G.; Feligini, M.; Mancusi, A.; D'Avino, A.; Coletta, A.; Berardino, D. D.; Ramunno, L. *Italian Mediterranean River Buffalo CSN2 Gene Structure and Promoter Analysis*. *Italian Journal of Animal Science* **2009**, *8* (sup2), 57–59. https://doi.org/10.4081/ijas.2009.s2.57.
36. Cosenza, G.; Gallo, D.; Auzino, B.; Gaspa, G.; Pauciullo, A. Complete CSN1S2 Characterization, Novel Allele Identification and Association With Milk Fatty Acid Composition in River Buffalo. *Front. Genet.* **2021**, *11*, 622494. https://doi.org/10.3389/fgene.2020.622494.
37. Cosenza, G.; Iannaccone, M.; Pico, B. A.; Ramunno, L.; Capparelli, R. The SNP g.1311T>C Associated with the Absence of *β* -casein in Goat Milk Influences *CSN 2* Promoter Activity. *Animal Genetics* **2016**, *47* (5), 615–617. https://doi.org/10.1111/age.12443.
38. Cosenza, G.; Illario, R.; Rando, A.; Di Gregorio, P.; Masina, P.; Ramunno, L. Molecular Characterization of the Goat *CSN1S1* ^{01} Allele. *Journal of Dairy Research* **2003**, *70* (2), 237–240. https://doi.org/10.1017/S0022029903006101.
39. Cosenza, G.; Pauciullo, A.; Coletta, A.; Di Francia, A.; Feligini, M.; Gallo, D.; Di Berardino, D.; Ramunno, L. Short Communication: Translational Efficiency of Casein Transcripts in Mediterranean River Buffalo. *Journal of Dairy Science* **2011**, *94* (11), 5691–5694. https://doi.org/10.3168/jds.2010-4086.
40. Cosenza, G.; Pauciullo, A.; Colimoro, L.; Mancusi, A.; Rando, A.; Di Berardino, D.; Ramunno, L. An SNP in the Goat CSN2 Promoter Region Is Associated with the Absence of β -casein in Milk. Animal Genetics 2007, 38 (6), 655–658. https://doi.org/10.1111/j.1365-2052.2007.01649.x.
41. Cosenza, G.; Pauciullo, A.; Gallo, D.; Berardino, D. D.; Ramunno, L. A *Ssp* I PCR-RFLP Detecting a Silent Allele at the Goat *CSN2* Locus. *Journal of Dairy Research* **2005**, *72* (4), 456–459. https://doi.org/10.1017/S0022029905001342.
42. Cosenza, G.; Pauciullo, A.; Gallo, D.; Colimoro, L.; D'Avino, A.; Mancusi, A.; Ramunno, L. Genotyping at the CSN1S1 Locus by PCR-RFLP and AS-PCR in a Neapolitan Goat Population. *Small Ruminant Research* **2008**, *74* (1–3), 84–90. https://doi.org/10.1016/j.smallrumres.2007.03.010.

43. Cosenza, G.; Pauciullo, A.; Macciotta, N. P. P.; Apicella, E.; Steri, R.; La Battaglia, A.; Jemma, L.; Coletta, A.; Di Berardino, D.; Ramunno, L. Mediterranean River Buffalo CSN1S1 Gene: Search for Polymorphisms and Association Studies. *Anim. Prod. Sci.* **2015**, *55* (5), 654. https://doi.org/10.1071/AN13438.
44. Cosenza, G.; Pauciullo, A.; Mancusi, A.; D'Avino, A.; Colimoro, L.; Di Berardino, D.; Ramunno, L. Identification of Single Nucleotide Polymorphisms in the Nicastrese Goat and Sardinia Sheep Mannose-Binding Lectin. *Biochem Genet* **2012**, *50* (1–2), 73–83. https://doi.org/10.1007/s10528-011-9461-5.
45. Cunsolo, V.; Muccilli, V.; Saletti, R.; Marletta, D.; Foti, S. Detection and Characterization by High-performance Liquid Chromatography and Mass Spectrometry of Two Truncated Goat α_{s2}-caseins. *Rapid Comm Mass Spectrometry* **2006**, *20* (7), 1061–1070. https://doi.org/10.1002/rcm.2415.
46. De Freitas, A. C.; De Camargo, G. M. F.; Stafuzza, N. B.; Aspilcueta-Borquis, R. R.; Venturini, G. C.; Dias, M. M.; Cardoso, D. F.; Tonhati, H. Genetic Association between SNPs in the DGAT1 Gene and Milk Production Traits in Murrah Buffaloes. *Trop Anim Health Prod* **2016**, *48* (7), 1421–1426. https://doi.org/10.1007/s11250-016-1110-x.
47. De Oliveira, L. S. M.; Alves, J. S.; Bastos, M. S.; Da Cruz, V. A. R.; Pinto, L. F. B.; Tonhati, H.; Costa, R. B.; De Camargo, G. M. F. Water Buffaloes (Bubalus Bubalis) Only Have A2A2 Genotype for Beta-Casein. *Trop Anim Health Prod* **2021**, *53* (1), 145. https://doi.org/10.1007/s11250-021-02591-9.
48. Dervishi, E.; Serrano, M.; Joy, M.; Sarto, P.; Somera, A.; González-Calvo, L.; Berzal-Herranz, B.; Molino, F.; Martinez-Royo, A.; Calvo, J. H. Structural Characterisation of the Acyl CoA: Diacylglycerol Acyltransferase 1 (DGAT1) Gene and Association Studies with Milk Traits in Assaf Sheep Breed. *Small Ruminant Research* **2015**, *131*, 78–84. https://doi.org/10.1016/j.smallrumres.2015.08.015.
49. Du, C.; Deng, T. X.; Zhou, Y.; Ghanem, N.; Hua, G. H. Bioinformatics Analysis of Candidate Genes for Milk Production Traits in Water Buffalo (Bubalus Bubalis). *Trop Anim Health Prod* **2020**, *52* (1), 63–69. https://doi.org/10.1007/s11250-019-01984-1.
50. Du, C.; Deng, T.; Zhou, Y.; Ye, T.; Zhou, Z.; Zhang, S.; Shao, B.; Wei, P.; Sun, H.; Khan, F. A.; Yang, L.; Hua, G. Systematic Analyses for Candidate Genes of Milk Production Traits in Water Buffalo (*Bubalus Bubalis*). *Animal Genetics* **2019**, *50* (3), 207–216. https://doi.org/10.1111/age.12739.
51. Erhardt, G., Jäger, S., Budelli, E., and Caroli, A. (2002). Genetic polymorphism of goat αS2-casein (CSN1S2) and evidence for a further allele. Milchwissenschaft 57, 137–140.
52. Farrell, H. M.; Jimenez-Flores, R.; Bleck, G. T.; Brown, E. M.; Butler, J. E.; Creamer, L. K.; Hicks, C. L.; Hollar, C. M.; Ng-Kwai-Hang, K. F.; Swaisgood, H. E. Nomenclature of the Proteins of Cows' Milk—Sixth Revision. *Journal of Dairy Science* **2004**, *87* (6), 1641–1674. https://doi.org/10.3168/jds.S0022-0302(04)73319-6.
53. Ferranti, P.; Malorni, A.; Nitti, G.; Laezza, P.; Pizzano, R.; Chianese, L.; Addeo, F. Primary Structure of Ovine α_{s1}-Caseins: Localization of Phosphorylation Sites and Characterization of Genetic Variants A, C and D. *Journal of Dairy Research* **1995**, *62* (2), 281–296. https://doi.org/10.1017/S0022029900030983
54. Galliano, F.; Saletti, R.; Cunsolo, V.; Foti, S.; Marletta, D.; Bordonaro, S.; D'Urso, G. Identification and Characterization of a New β-casein Variant in Goat Milk by High-performance Liquid Chromatography with Electrospray Ionization Mass Spectrometry and Matrix-assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Comm Mass Spectrometry* **2004**, *18* (17), 1972–1982. https://doi.org/10.1002/rcm.1575.
55. Gallinat, J. L.; Qanbari, S.; Drögemüller, C.; Pimentel, E. C. G.; Thaller, G.; Tetens, J. DNA-Based Identification of Novel Bovine Casein Gene Variants. *Journal of Dairy Science* **2013**, *96* (1), 699–709. https://doi.org/10.3168/jds.2012-5908.
56. Giambra IJ, Erhardt G. Molecular genetic characterization of ovine CSN1S2 variants C and D reveal further important variability within CSN1S2. Anim Genet. 2012 Oct;43(5):642-5. doi: 10.1111/j.1365-2052.2011.02299.x. Epub 2011 Dec 7. PMID: 22497628.
57. Grisart, B.; Coppieters, W.; Farnir, F.; Karim, L.; Ford, C.; Berzi, P.; Cambisano, N.; Mni, M.; Reid, S.; Simon, P.; Spelman, R.; Georges, M.; Snell, R. Positional Candidate Cloning of a QTL in Dairy Cattle: Identification of a Missense Mutation in the Bovine *DGAT1* Gene with Major Effect on Milk Yield and Composition. *Genome Res.* **2002**, *12* (2), 222–231. https://doi.org/10.1101/gr.224202.
58. Grisart, B.; Farnir, F.; Karim, L.; Cambisano, N.; Kim, J.-J.; Kvasz, A.; Mni, M.; Simon, P.; Frère, J.-M.; Coppieters, W.; Georges, M. Genetic and Functional Confirmation of the Causality of the *DGAT1 K232A* Quantitative Trait Nucleotide in Affecting Milk Yield and Composition. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (8), 2398–2403. https://doi.org/10.1073/pnas.0308518100.
59. Grisart, J. M.; Van Der Linden, M. Conscious and Automatic Uses of Memory in Chronic Pain Patients. *Pain* **2001**, *94* (3), 305–313. https://doi.org/10.1016/S0304-3959(01)00366-9.
- Groenen, M.A.M. and Van der Poel, J.J.: Regulation of expression of milk protein genes: a review. Livest. Prod. Sci. 38 (1994) 61-78.
60. Grosclaude, F.; Mahé, M.-F.; Brignon, G.; Di Stasio, L.; Jeunet, R. A Mendelian Polymorphism Underlying Quantitative Variations of Goat As1-Casein. *Genet Sel Evol* **1987**, *19* (4), 399. https://doi.org/10.1186/1297-9686-19-4-399.
61. Gu, M.; Cosenza, G.; Nicolae, I.; Bota, A.; Guo, Y.; Di Stasio, L.; Pauciullo, A. Transcript Analysis at DGAT1 Reveals Different mRNA Profiles in River Buffaloes with Extreme Phenotypes for Milk Fat. *Journal of Dairy Science* **2017**, *100* (10), 8265–8276. https://doi.org/10.3168/jds.2017-12771.

62. Haagsman, H. P.; De Haas, C. G. M.; Geelen, M. J. H.; Van Golde, L. M. G. Regulation of Triacylglycerol Synthesis in the Liver a Decrease in Diacylglycerol Acyltransferase Activity after Treatment of Isolated Rat Hepatocytes with Glucagon. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1981**, *664* (1), 74–81. https://doi.org/10.1016/0005-2760(81)90029-1.
63. Hassanin AA, Osman A, Atallah OO, El-Saadony MT, Abdelnour SA, Taha HSA, Awad MF, Elkashef H, Ahmed AE, Abd El-Rahim I, Mohamed A, Eldomiaty AS. Phylogenetic comparative analysis: Chemical and biological features of caseins (alpha-S-1, alpha-S-2, beta- and kappa-) in domestic dairy animals. Front Vet Sci. 2022 Sep 15;9:952319. doi: 10.3389/fvets.2022.952319. PMID: 36187819; PMCID: PMC9519386.
64. Homa, S. T.; Racowsky, C.; McGaughey, R. W. Lipid Analysis of Immature Pig Oocytes. *Reproduction* **1986**, *77* (2), 425–434. https://doi.org/10.1530/jrf.0.0770425.
65. Ibeagha-Awemu, E. M.; Kgwatalala, P.; Zhao, X. A Critical Analysis of Production-Associated DNA Polymorphisms in the Genes of Cattle, Goat, Sheep, and Pig. Mamm Genome 2008, 19 (9), 591–617. https://doi.org/10.1007/s00335-008-9141-x.
66. Ilie, D. E.; Mizeranschi, A. E.; Mihali, C. V.; Neamț, R. I.; Goilean, G. V.; Georgescu, O. I.; Zaharie, D.; Carabaș, M.; Huțu, I. Genome-Wide Association Studies for Milk Somatic Cell Score in Romanian Dairy Cattle. *Genes* **2021**, *12* (10), 1495. https://doi.org/10.3390/genes12101495.
67. Iung, L. H. S.; Petrini, J.; Ramírez-Díaz, J.; Salvian, M.; Rovadoscki, G. A.; Pilonetto, F.; Dauria, B. D.; Machado, P. F.; Coutinho, L. L.; Wiggans, G. R.; Mourão, G. B. Genome-Wide Association Study for Milk Production Traits in a Brazilian Holstein Population. *Journal of Dairy Science* **2019**, *102* (6), 5305–5314. https://doi.org/10.3168/jds.2018-14811.
68. Jansà Pérez M, Leroux C, Sànchez Bonastre A and Martin P. Occurrence of a LINE sequence in the 3' UTR of the goat as1-casein E-encoding allele associated with reduced protein synthesis level. Gene. 147 (1994) 179-187
69. Kamiński, S.; Cieślińska, A.; Kostyra, E. Polymorphism of Bovine Beta-Casein and Its Potential Effect on Human Health. *J Appl Genet* **2007**, *48* (3), 189–198. https://doi.org/10.1007/BF03195213.
70. Khan, K.; Suhail, S. M.; Khan, R.; Ahmed, I.; Khan, F. A.; Khan, M. J. Genetic Polymorphism of B-Casein Gene and Its Association with Milk Production and Composition in Azi-Kheli Buffalo. *Trop Anim Health Prod* **2023**, *55* (2), 94. https://doi.org/10.1007/s11250-023-03511-9.
71. King JWB (1966) The caseins of sheep's milk. In: Polymorphismes Biochimiques des AnimauxParis: INRA Proceedings of 10th European Conference on Animal Blood Groups and Biochemical Polymorphism, pp 427–431
72. Kühn, C.; Thaller, G.; Winter, A.; Bininda-Emonds, O. R. P.; Kaupe, B.; Erhardt, G.; Bennewitz, J.; Schwerin, M.; Fries, R. Evidence for Multiple Alleles at the DGAT1 Locus Better Explains a Quantitative Trait Locus With Major Effect on Milk Fat Content in Cattle. *Genetics* **2004**, *167* (4), 1873–1881. https://doi.org/10.1534/genetics.103.022749.
73. Lagonigro, R.; Pietrola, E.; D'Andrea, M.; Veltri, C.; Pilla, F. Molecular Genetic Characterization of the Goat s_2-casein *E* Allele. *Animal Genetics* **2001**, *32* (6), 391–393. https://doi.org/10.1046/j.1365-2052.2001.0781c.x.
74. Lajnaf, R.; Feki, S.; Attia, H.; Ayadi, M. A.; Masmoudi, H. Characteristics of Cow Milk Proteins and the Effect of Processing on Their Allergenicity. In *Milk Protein*; Chaiyabutr, N., Ed.; IntechOpen: Rijeka, 2022. https://doi.org/10.5772/intechopen.102494.
75. Lassner M. Transgenic oilseed crops: a transition from basic research to product development. Lipid Technol. 1997;9(5):9.
76. Lehner R, Kuksis A. Biosynthesis of triacylglycerols. *Prog. Lipid Res.* 1996;35:169–201.
77. Leroux C, Mazure N, Martin P. Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat alpha s1-casein transcripts. Structural organization of the relevant gene. J Biol Chem. 1992 Mar 25;267(9):6147-57. PMID: 1372900.
78. Li, J.; Liu, S.; Li, Z.; Zhang, S.; Hua, G.; Salzano, A.; Campanile, G.; Gasparrini, B.; Liang, A.; Yang, L. *DGAT1* Polymorphism in Riverine Buffalo, Swamp Buffalo and Crossbred Buffalo. *Journal of Dairy Research* **2018**, *85* (4), 412–415. https://doi.org/10.1017/S0022029918000468.
79. Lillie, B. N.; Brooks, A. S.; Keirstead, N. D.; Hayes, M. A. Comparative Genetics and Innate Immune Functions of Collagenous Lectins in Animals. *Veterinary Immunology and Immunopathology* **2005**, *108* (1–2), 97–110. https://doi.org/10.1016/j.vetimm.2005.07.001.
80. Lillie, B.; Hammermueller, J.; Macinnes, J.; Jacques, M.; Hayes, M. Porcine Mannan-Binding Lectin A Binds to Actinobacillus Suis and Haemophilus Parasuis. *Developmental & Comparative Immunology* **2006**, *30* (10), 954–965. https://doi.org/10.1016/j.dci.2005.12.008.
81. Liu, J.; Ju, Z.; Li, Q.; Huang, J.; Li, R.; Li, J.; Ma, L.; Zhong, J.; Wang, C. Mannose-Binding Lectin 1 Haplotypes Influence Serum MBL-A Concentration, Complement Activity, and Milk Production Traits in Chinese Holstein Cattle. *Immunogenetics* **2011**, *63* (11), 727–742. https://doi.org/10.1007/s00251-011-0548-2.
82. Mahè M.F., Grosclaude F. (1993). Polymorphism of β-casein in the Creole goat of Guadeloupe: evidence for a null allele Génétique, Seléction et Evolution 21, 127-129.
83. Mangia, N. P.; Saliba, L.; Zoumpopoulou, G.; Chessa, S.; Anastasiou, R.; Karayiannis, I.; Sgouras, D.; Tsakalidou, E.; Nudda, A. Goat Milk with Different Alpha-S1 Casein Genotype (CSN1S1) Fermented by Selected Lactobacillus Paracasei as Potential Functional Food. *Fermentation* **2019**, *5* (3), 55. https://doi.org/10.3390/fermentation5030055.

84. Martin, P., Bianchi, L., Cebo, C., Miranda, G., Mcsweeney, P., and Fox, P. (2013). "Genetic polymorphism of milk proteins," in Advanced dairy chemistry (Boston: Springer).
85. Martin, P., Ollivier-Bousquet, M., Grosclaude, F., 1999. Genetic polymorphism of caseins: a tool to investigate casein micelle organization. Int. Dairy J. 9, 163–170
86. Martin, P.; Palhière, I.; Maroteau, C.; Bardou, P.; Canale-Tabet, K.; Sarry, J.; Woloszyn, F.; Bertrand-Michel, J.; Racke, I.; Besir, H.; Rupp, R.; Tosser-Klopp, G. A Genome Scan for Milk Production Traits in Dairy Goats Reveals Two New Mutations in Dgat1 Reducing Milk Fat Content. *Sci Rep* **2017**, *7* (1), 1872. https://doi.org/10.1038/s41598-017-02052-0.
87. Massender, E.; Oliveira, H. R.; Brito, L. F.; Maignel, L.; Jafarikia, M.; Baes, C. F.; Sullivan, B.; Schenkel, F. S. Genome-Wide Association Study for Milk Production and Conformation Traits in Canadian Alpine and Saanen Dairy Goats. *Journal of Dairy Science* **2023**, *106* (2), 1168–1189. https://doi.org/10.3168/jds.2022-22223.
88. Meng, L., Chen, L., Miao, Y., Zhang, Y., Zhao, S., Huo, J., Li, W., 2013. Polymorphism in the exon 17 of *DGAT1* gene and its association with milk composition traits in buffalo (*Bubalus bubalis*). Journal of Yunnan Agricultural University, 2, 190–195., **2013**, 28(2): 190-195. DOI: 10.3969/j.issn.1004—390X(n).2013.02.008
89. Miglior F, Fleming A, Malchiodi F, Brito LF, Martin P, Baes CF. A 100-Year Review: Identification and genetic selection of economically important traits in dairy cattle. J Dairy Sci. **2017** Dec;100(12):10251-10271. doi: 10.3168/jds.2017-12968. PMID: 29153164.
90. Moatsou, G.; Mollé, D.; Moschopoulou, E.; Valérie, G. Study of Caprine β-Casein Using Reversed-Phase High-Performance Liquid Chromatography and Mass Spectroscopy: Identification of a New Genetic Variant. *Protein J* **2007**, *26* (8), 562–568. https://doi.org/10.1007/s10930-007-9098-8.
91. Nepomuceno, R. R.; Henschen-Edman, A. H.; Burgess, W. H.; Tenner, A. J. cDNA Cloning and Primary Structure Analysis of C1qRP, the Human C1q/MBL/SPA Receptor That Mediates Enhanced Phagocytosis In Vitro. *Immunity* **1997**, *6* (2), 119–129. https://doi.org/10.1016/S1074-7613(00)80419-7.
92. Neveu, C.; Mollé, D.; Moreno, J.; Martin, P.; Léonil, J. Heterogeneity of Caprine Beta-Casein Elucidated by RP-HPLC/MS: Genetic Variants and Phosphorylations. *J Protein Chem* **2002**, *21* (8), 557–567. https://doi.org/10.1023/A:1022433823559.
93. Nicolai, M. A.; Garro, G.; Caira, S.; Mauriello, R.; Quarto, M.; De Pascale, S.; Chianese, L.Occurrence of Quantitative Genetic Polymorphism at the Caprine β-CN Locus, as Determined by a Proteomic Approach. *International Dairy Journal* **2021**, *112*, 104855. https://doi.org/10.1016/j.idairyj.2020.104855.
94. Nonneman, D.; Rohrer, G. A. Linkage Mapping of Porcine *DGAT1* to a Region of Chromosome 4 That Contains QTL for Growth and Fatness [1,2,3]. *Animal Genetics* **2002**, *33* (6), 472–473. https://doi.org/10.1046/j.1365-2052.2002.00938_5.x.
95. Nuomin; Nguyen, Q. D.; Aodaohu; Nishino, N. Frequency of β-Casein Gene Polymorphisms in Jersey Cows in Western Japan. *Animals* **2022**, *12* (16), 2076. https://doi.org/10.3390/ani12162076.
96. Persuy, M.; Printz, C.; Medrano, J. F.; Mercier, J. A Single Nucleotide Deletion Resulting in a Premature Stop Codon Is Associated with Marked Reduction of Transcripts from a Goat B-casein Null Allele. *Animal Genetics* **1999**, *30* (6), 444–451. https://doi.org/10.1046/j.1365-2052.1999.00547.x.
97. Picariello, G.; Rignanese, D.; Chessa, S.; Ceriotti, G.; Trani, A.; Caroli, A.; Di Luccia, A. Characterization and Genetic Study of the Ovine α S2 -Casein (CSN1S2) Allele B. *Protein J* **2009**, *28* (7–8), 333–340. https://doi.org/10.1007/s10930-009-9198-8.
98. Pineda P.S., Delos Santos J.D., Flores E.B. GENETIC POLYMORPHISM OF β-CASEIN EXON 7 IN BUFFALOES (Bubalus bubalis). Philipp J Vet Anim Sci **2019** 45(3):197-202
99. Rahmatalla, S. A.; Arends, D.; Brockmann, G. A. Review: Genetic and Protein Variants of Milk Caseins in Goats. *Frontiers in Genetics* **2022**, *13*.
100. Rahmatalla, S. A.; Arends, D.; Said Ahmed, A.; Hassan, L. M. A.; Krebs, S.; Reissmann, M.; Brockmann, G. A. Capture Sequencing to Explore and Map Rare Casein Variants in Goats. *Front. Genet.* **2021**, *12*, 620253. https://doi.org/10.3389/fgene.2021.620253.
101. Ramesha, K P; Rao, Akhila; Basavaraju, M; Alex, Rani; Kataktalware, M A; Jeyakumar, S; Varalakshmi, S. Genetic variants of β-casein in cattle and buffalo breeding bulls in Karnataka state of India. IJBT Vol.15(2) [April 2016] 178-181
102. Ramunno L, Cosenza G, Crepaldi P, Pilla F. Biologia molecolare e miglioramento genetico delle caratteristiche quali-quantitative del latte ovi-caprino. Supplemento Large Animal Review 2008; 14:117-121
103. Ramunno L, Pauciullo A, Mancusi A, Cosenza G, Mariani P, Malacarne M. Influenza del polimorfismo genetico delle caseine calcio-sensibili su caratteristiche strutturali, nutrizionali, attitudine casearia e proprietà ipoallergeniche del latte di capra. Scienza e tecnica lattiero-casearia, *58* (4), 257-271, 2007.
104. Ramunno L., Mariani P., Pappalardo M., Rando A., Capuano M., Di Gregorio P., Cosenza G. (1995).Un gene a effetto maggiore sul contenuto di caseina β nel latte di capra. XI Congr. ASPA, 185-186.
105. Ramunno, L.; Cosenza, G.; Pappalardo, M.; Longobardi, E.; Gallo, D.; Pastore, N.; Di Gregorio, P.; Rando, A. Characterization of Two New Alleles at the Goat *CSN1S2* Locus. *Animal Genetics* **2001a**, *32* (5), 264–268. https://doi.org/10.1046/j.1365-2052.2001.00786.x.

106. Ramunno, L.; Cosenza, G.; Rando, A.; Pauciullo, A.; Illario, R.; Gallo, D.; Di Berardino, D.; Masina, P. Comparative Analysis of Gene Sequence of Goat CSN1S1 F and N Alleles and Characterization of CSN1S1 Transcript Variants in Mammary Gland. *Gene* **2005**, *345* (2), 289–299. https://doi.org/10.1016/j.gene.2004.12.003.
107. Ramunno, L.; Longobardi, E.; Pappalardo, M.; Rando, A.; Di Gregorio, P.; Cosenza, G.; Mariani, P.; Pastore, N.; Masina, P. An Allele Associated with a Non-detectable Amount of As $_2$ Casein in Goat Milk. *Animal Genetics* **2001b**, *32* (1), 19–26. https://doi.org/10.1046/j.1365-2052.2001.00710.x.
108. Rando A., Pappalardo M., Capuano M., Di Gregorio P., Ramunno L. (1996). Two mutations might be responsible for the absence of b-casein in goat milk. Anim. Genet. 27, 31.
109. Rando, A.; Di Gregorio, P.; Ramunno, L.; Mariani, P.; Fiorella, A.; Senese, C.; Marletta, D.; Masina, P. Characterization of the CSN1AG Allele of the Bovine Alpha S1-Casein Locus by the Insertion of a Relict of a Long Interspersed Element. *J Dairy Sci* **1998**, *81* (6), 1735–1742. https://doi.org/10.3168/jds.S0022-0302(98)75741-8.
110. Richardson, B. C.; Creamer, L. K. Comparative Micelle Structure. III. The Isolation and Chemical Characterization of Caprine And. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1974**, *365* (1), 133–137. https://doi.org/10.1016/0005-2795(74)90257-8.
111. Roberts, B.; DiTullio, P.; Vitale, J.; Hehir, K.; Gordon, K. Cloning of the Goat β-Casein-Encoding Gene and Expression in Transgenic Mice. *Gene* **1992**, *121* (2), 255–262. https://doi.org/10.1016/0378-1119(92)90129-D.
112. Roin, N. R.; Larsen, L. B.; Comi, I.; Devold, T. G.; Eliassen, T. I.; Inglingstad, R. A.; Vegarud, G. E.; Poulsen, N. A. Identification of Rare Genetic Variants of the αS-Caseins in Milk from Native Norwegian Dairy Breeds and Comparison of Protein Composition with Milk from High-Yielding Norwegian Red Cows. *Journal of Dairy Science* **2022**, *105* (2), 1014–1027. https://doi.org/10.3168/jds.2021-20455.
113. Sacchi, P.; Chessa, S.; Budelli, E.; Bolla, P.; Ceriotti, G.; Soglia, D.; Rasero, R.; Cauvin, E.; Caroli, A. Casein Haplotype Structure in Five Italian Goat Breeds. *Journal of Dairy Science* **2005**, *88* (4), 1561–1568. https://doi.org/10.3168/jds.S0022-0302(05)72825-3.
114. Scatà, M. C.; Napolitano, F.; Casu, S.; Carta, A.; De Matteis, G.; Signorelli, F.; Annicchiarico, G.; Catillo, G.; Moioli, B. Ovine *Acyl CoA :Diacylglycerol Acyltransferase 1* – Molecular Characterization, Polymorphisms and Association with Milk Traits. *Animal Genetics* **2009**, *40* (5), 737–742. https://doi.org/10.1111/j.1365-2052.2009.01909.x.
115. Sebastiani, C.; Arcangeli, C.; Ciullo, M.; Torricelli, M.; Cinti, G.; Fisichella, S.; Biagetti, M. Frequencies Evaluation of β-Casein Gene Polymorphisms in Dairy Cows Reared in Central Italy. *Animals* **2020**, *10* (2), 252. https://doi.org/10.3390/ani10020252.
116. Secchiari P.L., Carnier P., Priolo A., Mele M. Basi genetiche e fisiologiche della qualità degli alimenti di origine animale. *Ital. J. Agron. / Riv. Agron., 2009, 1 Suppl.:81-102*
117. Selionova, M.; Trukhachev, V.; Aibazov, M.; Sermyagin, A.; Belous, A.; Gladkikh, M.; Zinovieva, N. Genome-Wide Association Study of Milk Composition in Karachai Goats. *Animals* **2024**, *14* (2), 327. https://doi.org/10.3390/ani14020327.
118. Selvaggi, M.; Laudadio, V.; Dario, C.; Tufarelli, V. Investigating the Genetic Polymorphism of Sheep Milk Proteins: A Useful Tool for Dairy Production. *J Sci Food Agric* **2014b**, *94* (15), 3090–3099. https://doi.org/10.1002/jsfa.6750.
119. Selvaggi, M.; Laudadio, V.; Dario, C.; Tufarelli, V. Major Proteins in Goat Milk: An Updated Overview on Genetic Variability. *Mol Biol Rep* **2014a**, *41* (2), 1035–1048. https://doi.org/10.1007/s11033-013-2949-9.
120. Shergojry, S. A.; Verma, A.; Ghani, M.; Gupta, I. D.; Mir, N. A. Identification of Genetic Polymorphism of the MBL2 Gene and Its Association with Clinical Mastitis in Murrah Buffaloes. *J Genet* **2023**, *102* (1), 21. https://doi.org/10.1007/s12041-023-01419-9.
121. Shi, L.; Takahashi, K.; Dundee, J.; Shahroor-Karni, S.; Thiel, S.; Jensenius, J. C.; Gad, F.; Hamblin, M. R.; Sastry, K. N.; Ezekowitz, R. A. B. Mannose-Binding Lectin-Deficient Mice Are Susceptible to Infection with *Staphylococcus Aureus*. *The Journal of Experimental Medicine* **2004**, *199* (10), 1379–1390. https://doi.org/10.1084/jem.20032207.
122. Singh, U.; Deb, R.; Alyethodi, R. R.; Alex, R.; Kumar, S.; Chakraborty, S.; Dhama, K.; Sharma, A. Molecular Markers and Their Applications in Cattle Genetic Research: A Review. Biomarkers and Genomic Medicine 2014, 6 (2), 49–58. https://doi.org/10.1016/j.bgm.2014.03.001.
123. Smith, Lisa C. and Lawrence James Haddad. Overcoming child malnutrition in developing countries: past achievements and future choices . vol. 30. International Food Policy Res Inst, 2000.
124. Spelman, R. J.; Ford, C. A.; McElhinney, P.; Gregory, G. C.; Snell, R. G. Characterization of the DGAT1 Gene in the New Zealand Dairy Population. *Journal of Dairy Science* **2002**, *85* (12), 3514–3517. https://doi.org/10.3168/jds.S0022-0302(02)74440-8.
125. Sukla, S.; Bhattacharya, T. K.; Venkatachalapathy, R. T.; Kumar, P.; Sharma, A. Molecular Cloning and Characterization of Buffalo α $_{s1}$ -Casein Gene: Full Length Research Paper. *DNA Sequence* **2007**, *18* (5), 334–340. https://doi.org/10.1080/10425170701388974.
126. Sutera, A. M.; Moscarelli, A.; Mastrangelo, S.; Sardina, M. T.; Di Gerlando, R.; Portolano, B.; Tolone, M. Genome-Wide Association Study Identifies New Candidate Markers for Somatic Cells Score in a Local Dairy Sheep. *Front. Genet.* **2021**, *12*, 643531. https://doi.org/10.3389/fgene.2021.643531.

127. Swanton, E. M. S.; Saggerson, E. D. Glycerolipid Metabolizing Enzymes in Rat Ventricle and in Cardiac Myocytes. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1997**, *1346* (1), 93–102. https://doi.org/10.1016/S0005-2760(97)00024-6.
128. Takahashi, K.; Shi, L.; Gowda, L. D.; Ezekowitz, R. A. B. Relative Roles of Complement Factor 3 and Mannose-Binding Lectin in Host Defense against Infection. *Infect Immun* **2005**, *73* (12), 8188–8193. https://doi.org/10.1128/IAI.73.12.8188-8193.2005.
129. Thaller, G.; Krämer, W.; Winter, A.; Kaupe, B.; Erhardt, G.; Fries, R. Effects of DGAT1 Variants on Milk Production Traits in German Cattle Breeds1. *Journal of Animal Science* **2003**, *81* (8), 1911–1918. https://doi.org/10.2527/2003.8181911x.
130. Thiel, S.; Vorup-Jensen, T.; Stover, C. M.; Schwaeble, W.; Laursen, S. B.; Poulsen, K.; Willis, A. C.; Eggleton, P.; Hansen, S.; Holmskov, U.; Reid, K. B. M.; Jensenius, J. C. A Second Serine Protease Associated with Mannan-Binding Lectin That Activates Complement. *Nature* **1997**, *386* (6624), 506–510. https://doi.org/10.1038/386506a0.
131. Torres-Vazquez, J. A.; Vazquez Flores, F.; Montaldo, H. H.; Ulloa-Arvizu, R.; Valencia Posadas, M.; Gayosso Vazquez, A.; Alonso Morales, R. A. Genetic Polymorphism of the As1-Casein Locus in Five Populations of Goats from Mexico. *Electron. J. Biotechnol.* **2008**, *11* (3), 0–0. https://doi.org/10.2225/vol11-issue3-fulltext-11.
132. Turner, M. W. Mannose-Binding Lectin (MBL) in Health and Disease. *Immunobiology* **1998**, *199* (2), 327–339. https://doi.org/10.1016/S0171-2985(98)80037-5.
133. Van De Wetering, J. K.; Van Golde, L. M. G.; Batenburg, J. J. Collectins: Players of the Innate Immune System. *European Journal of Biochemistry* **2004**, *271* (7), 1229–1249. https://doi.org/10.1111/j.1432-1033.2004.04040.x.
134. Veltri, C., Lagonigro, R., Pietrolà, E., D'andrea, M., Pilla, F., and Chianese, L. (2000). "Molecular characterization of the goat αs2-casein E allele and its detection in goat breeds of Italy," in Proceedings 7th International Conference on Goats, Tour, France.
135. Wang, C.; Liu, M.; Li, Q.; Ju, Z.; Huang, J.; Li, J.; Wang, H.; Zhong, J. Three Novel Single-Nucleotide Polymorphisms of MBL1 Gene in Chinese Native Cattle and Their Associations with Milk Performance Traits. *Veterinary Immunology and Immunopathology* **2011**, *139* (2–4), 229–236. https://doi.org/10.1016/j.vetimm.2010.10.023.
136. Wang, M.; Song, H.; Zhu, X.; Xing, S.; Zhang, M.; Zhang, H.; Wang, X.; Yang, Z.; Ding, X.; Karrow, N. A.; König, S.; Mao, Y. *Toll-like Receptor* 4 Gene Polymorphisms Influence Milk Production Traits in Chinese Holstein Cows. *Journal of Dairy Research* **2018**, *85* (4), 407–411. https://doi.org/10.1017/S0022029918000535.
137. Wang, X.; Ju, Z.; Huang, J.; Hou, M.; Zhou, L.; Qi, C.; Zhang, Y.; Gao, Q.; Pan, Q.; Li, G.; Zhong, J.; Wang, C. The Relationship between the Variants of the Bovine MBL2 Gene and Milk Production Traits, Mastitis, Serum MBL-C Levels and Complement Activity. *Veterinary Immunology and Immunopathology* **2012**, *148* (3–4), 311–319. https://doi.org/10.1016/j.vetimm.2012.06.017.
138. Weller, J. I.; Golik, M.; Seroussi, E.; Ezra, E.; Ron, M. Population-Wide Analysis of a QTL Affecting Milk-Fat Production in the Israeli Holstein Population. *Journal of Dairy Science* **2003**, *86* (6), 2219–2227. https://doi.org/10.3168/jds.S0022-0302(03)73812-0.
139. Winter, A.; Krämer, W.; Werner, F. A. O.; Kollers, S.; Kata, S.; Durstewitz, G.; Buitkamp, J.; Womack, J. E.; Thaller, G.; Fries, R. Association of a Lysine-232/Alanine Polymorphism in a Bovine Gene Encoding Acyl-CoA:Diacylglycerol Acyltransferase (*DGAT1*) with Variation at a Quantitative Trait Locus for Milk Fat Content. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (14), 9300–9305. https://doi.org/10.1073/pnas.142293799.
140. Yuan, J.; Zhou, J.; Deng, X.; Hu, X.; Li, N. Molecular Cloning and Single Nucleotide Polymorphism Detection of Buffalo DGAT1 Gene. *Biochem Genet* **2007**, *45* (7–8), 611–621. https://doi.org/10.1007/s10528-007-9100-3.
141. Zhang, Y.; Fan, X.; Zhou, F.; Li, W.; Ouyang, Y.; Miao, Y. Polymorphisms of the CSN1S1 Gene and Its Protein Variants in River and Swamp Buffalo (Bubalus Bubalis). *PJZ* **2021**, *53* (4). https://doi.org/10.17582/journal.pjz/20191202121250.
142. Zhao, F.; Zhao, Z.; Yan, G.; Wang, D.; Ban, Q.; Yu, P.; Zhang, W.; Luo, Y. Polymorphisms in Mannose-Binding Lectin (MBL) Gene and Their Association with MBL Protein Levels in Serum in the Hu Sheep. *Veterinary Immunology and Immunopathology* **2011**, *140* (3–4), 297–302. https://doi.org/10.1016/j.vetimm.2010.12.009.
143. Zhao, Z. L.; Wang, C. F.; Li, Q. L.; Ju, Z. H.; Huang, J. M.; Li, J. B.; Zhong, J. F.; Zhang, J. B. Novel SNPs of the Mannan-Binding Lectin 2 Gene and Their Association with Production Traits in Chinese Holsteins. *Genet. Mol. Res.* **2012**, *11* (4), 3744–3754. https://doi.org/10.4238/2012.October.15.6.
144. Zicarelli, L.; Di Palo, R.; Napolano, R.; Tonhati, H.; De Carlo, E.; Gagliardi, R.; Di Luccia, A.; La Gatta, B. Influence of αS1-Casein and κ-Casein Polymorphism on the Curd Yield of Italian Mediterranean Buffalo (Bubalus Bubalis L.) Milk. *International Dairy Journal* **2020**, *100*, 104559. https://doi.org/10.1016/j.idairyj.2019.104559.

GENERAL AIMS

The main aim of this thesis has been the genetic characterization of central-southern Italy livestock populations, focusing on genes widely investigated for milk productions, in order to preserve them and ensure economic sustainability of farming.

These kinds of studies are essential for developing native breeds targeted conservation strategies that can preserve their genetic diversity and ensure their survival, contributing to biodiversity safeguard. They also support the sustainability of local communities by promoting traditional agricultural practices and preserving cultural ties to the land. In addition, studying the genetic basis of their adaptability to local conditions means they can play an important role in addressing global challenges such as climate change and food security. Our integrated approach aims to revolutionize breeding by wisely combining animal genetics with holistic management that considers the environment, animal welfare and overall sustainability.

To this purpose the Autochthonous Genetic Types (TGA), Bianca Monticellana, Capestrina, Grigia Ciociara (Fig. 10, 11, 12), Comisana and Pecora Bagnolese has been studied. Additionally, the Italian Mediterranean Buffalo has also been involved, although it is not a TGA, due to its economic role in the zoothecnics of central-southern Italy. This buffalo breed has unique distinctive traits that have been selected over years of genetic isolation and the lack of crossbreeding with other buffalo breeds. This led to its official recognition as a breed in 2000, making it imperative characterize it to enhance the valorization of its products.

The growing interest that institutions show towards these breeds is essentially expressed through the implementation of fundings aimed at enhancing and, consequently, deepening the genetic knowledge of these breeds. Infact, the studies here described have been funded by projects such as SAVEPEB, BIG, and AGRITECH, each of which will be briefly outlined below.

The acronym SAVEPEB identifies the project "Environmental preservation and economic enhancement of the Bagnolese Sheep" financed through the 2014-2020 Rural Development Program – Campania Region. SAVEPEB purpose was improved the competitiveness of the short supply chain of Pecorino Bagnolese through a quality regime that involved both the breeding of the autochthonous Pecora Bagnolese breed and the Pecorino cheese production process.

The acronym BIG identifies the project “Italian Mediterranean Buffalo - innovative technologies for genetic improvement”, financed through National Rural Development Program - Submeasure 10.2 - Ministry of Agriculture, Food Sovereignty and Forestry spanning three years (2020 to present). The BIG aim is to genetically enhance innovative traits related to resistance/resilience to diseases (brucellosis, tuberculosis, paratuberculosis, transverse hemimelia), environmental sustainability (feed efficiency, atmospheric emissions), production quality (cheesemaking), and the promotion of meat

production as a secondary product (e.g., muscularity and BCS) in the Italian Mediterranean Buffalo breed. The project also focuses on maintaining genetic variability by controlling inbreeding and establishing a biobank for male semen and embryos.

Finally, the acronym AGRITECH identifies the National Center for Agricultural Technology that was funded by the National Recovery and Resilience Plan - Ministry of University and Research for a project based on the use of enabling technologies for the sustainable development of agri-food productions. Its aim is to promote adaptation to climate change, reduce environmental impact in the agrifood sector, develop marginal areas, and enhance the safety, traceability, and typical characteristics of supply chains.

To give farmers valuable tools and information to improve their productions by leveraging the distinctive traits arising from the genetics of the animals, environmental peculiarities, and breeding practices, we investigated:

Chapter 1: the *CSN1S1*, *CSN2*, *CSN1S2* genes in the autochthonous breeds Bianca Monticellana, Capestrina, and Grigia Ciociara;

Chapter 2: the *DGAT1* gene in the Pecora Bagnolese;

Chapter 3: *MBL2* gene in Italian Mediterranean Buffalo

Chapter 4: the *MBL2* gene in Comisana sheep

In all the above-mentioned studies the main objective was to investigate the genetic structure of the breeds concerned at specific loci (*CSN1S1*, *CSN2*, *CSN1S2*, *DGAT1* and *MBL2*) and to study their association with production. The purpose was to apply this knowledge within the production chain in order to optimize the management practices and to develop well suited valorization plans of the derived products.

The results of these studies will not only contribute to the preservation of local animal populations but will also enhance the economic sustainability of farming in the specific context of central-southern Italy. Active collaboration with farmers and the dissemination of the obtained results are key elements for the success and positive impact of our research on the local zootechnical community.

Fig. 10. Bianca Monticellana goat breed female (left) and male (right) goats.

Fig. 11. Capestrina goat breed female (left) and male (right) goats.

Fig. 12. Grigia Ciociara goat breed female (left) and male (right) goats.

CHAPTER 1

A New AS-PCR Method to Detect *CSN2*^{01} Allele, Genotyping at Ca-Sensitive Caseins Loci and Milk Traits Association Studies in Autochthonous Lazio Goats

Gianfranco Cosenza [1],†, Sara Albarella [2],† , Emanuele D'Anza [2], Alessandra Iannuzzi [3]* , Maria Selvaggi [4]* Mariagiulia Pugliano [2], Tiziana Galli [5], Giorgio Saralli [5], Francesca Ciotola [2] and Vincenzo Peretti [2]

1 Department of Agriculture, University of Napoli Federico II, 80055 Portici, Italy

2 Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Via Delpino 1, 80137 Naples, Italy

3 National Research Council (CNR), Institute of Animal Production System in Mediterranean Environment (ISPAAM), Piazzale E. Fermi, 1, 8055 Portici, Italy

4 Department of Soil, Plant and Food Sciences, University of Bari Aldo Moro, 70126 Bari, Italy

5 Istituto Zooprofilattico Sperimentale Lazio e Toscana "M. Aleandri" (IZSLT), UOT Lazio Sud, Str. Congiunte Destre, 04100 Latina, Italy

* Correspondence: alessandra.iannuzzi@cnr.it (A.I.); maria.selvaggi@uniba.it (M.S.)

† These authors contributed equally to this work.

Simple Summary: Identifying new alleles and mutations at calcium-sensitive casein *loci* requires continuous updating of genotyping protocols useful for association study. This study reports (1) a new, more efficient, specific PCR-based genotyping protocol to detect the *CSN2*^{01} allele in goats, (2) the genetic characterization at the *CSN1S1*, *CSN2*, and *CSN1S2 loci* in three endangered goat breeds reared in the Lazio Region (Central Italy) and, (3) the association of the genotypes observed in the studied animals with parameters that might affect the milk's traits. As regards the *CSN1S1*, *CSN2*, and *CSN1S2 loci*, no animals were found to be carriers of the *CSN1S1*^{01}, *CSN1S1*^{E}, *CSN2*^{01}, *CSN1S2*^{D}, and *CSN1S2*^{0} alleles; instead, for the *CSN1S1 locus*, a high frequency of alleles associated to a low (*CSN1S1*^{F}) and high (*CSN1S1*^{A*,B*}) content of the αs1 casein (αs1-Cn) content in milk, with *CSN1S1*^{F}≥ *CSN1S1*^{B*} ≥ *CSN1S1*^{A*} being observed. An association between the different genotypes at the *CSN1S1 locus* and some milk traits, namely the fat and protein yielded and the fat, protein, solids-not-fat, and casein percentages without an effect on the milk yield, was observed.

Abstract: Calcium-sensitive caseins are the main protein component of milk. In the goat, they are encoded by three genes (*CSN1S1*, *CSN2*, and *CSN1S2*) located on chromosome 6. A high number of alleles has been discovered for these genes in the goat species, responsible for changes in the milk's qualitative and quantitative characteristics. This study aimed to develop an Allele-Specific PCR (AS-PCR), which allowed us to unequivocally detect goat carriers of the *CSN2*^{01} allele. Subsequently, the calcium-sensitive casein *loci* genotype was investigated in three native goat breeds of the Lazio Region (Bianca Monticellana, Capestrina, and Ciociara Grigia). No individuals were carriers of the *CSN1S1*^{01}, *CSN1S1*^{E}, *CSN2*^{01} , *CSN1S2*^{D} , and *CSN1S2*^{0} alleles, while a high frequency of the

alleles *CSN1S1*^{F} and *CSN1S1*^{A*,B*} was observed. Association analyses between the different genotypes at the *CSN1S1 locus* and some milk traits, namely the fat and protein yielded and the fat, protein, solids-not-fat, and casein percentages without an effect on the milk yield, were observed.

Keywords: calcium-sensitive caseins; genotyping; milk traits; autochthonous goat breeds

1. Introduction

Recent achievements in molecular genetics provide the opportunity to investigate genomic regions that directly or indirectly influence animal production. Moreover, special attention is given to the enhancement and protection of native breeds of different live- stock species, characterized by peculiar allele and haplotype combinations that make their production unique [1,2]. Among ruminants, the goat species shows the most significant genetic variability, especially at the *CSN1S1*, *CSN2,* and *CSN1S2 loci* that code for the three calcium-sensitive caseins: αs1-Cn, β-Cn (β casein), and αs2-Cn (αs2 casein), respectively. The *CSN1S1* gene is the most investigated and characterized by the highest level of poly- morphism. The number of alleles identified at this *locus* has increased significantly over the years. To date, there are at least 22 variants associated with qualitative and quantitative differences in the content of αs1-Cn and, thus, clustered as: 'strong' (A, A2, A3, A′, B′, B1, B2, B3, B4, C, H, L, and M: approximately 3.5 to 4.2 g/L per allele), 'intermediate' (E, I, D1: ~1.1 to 1.6 g/L per allele), 'weak' (D, F, and G: ~0.45 to 0.6 g/L per allele), and 'null' alleles (01, 02, and N: 0.0 g/L per allele or trace) [2–7]. Relatively recently, two new "null" and F- like αs1-casein variants have been detected only in an indigenous Norwegian goat breed [8]. Regarding the *CSN2* gene, at least ten alleles have been identified at this *locus*: the *CSN2* alleles, such as A, A1, B, C, C1, D, E, and F alleles, which are associated with a normal (strong) amount of α-casein in goat milk (about 5–6 g/L per allele), and the *CSN2* 0 and 01 alleles, which are associated with a non-detectable amount of this protein in milk [9,10]. Recently, two new β-Cn variants, C2 and F1, less expressed than the most common, were found [11]. Studies on the *CSN1S2 locus* led to the identification of at least eight alleles related to three different quantities of αs2-Cn in milk, clustered as: normal (*CSN1S2* A, B, C, E, F, and G: ~2.5 g/L per allele), intermediate (D: ~1.5 g/L per allele), or null (0: no detectable amount of this casein) [12–14]. Recently, four new non-defective αs2-Cn variants named *CSN1S2* H, I, J, and K, according to the existing alphabetical order for this protein, were detected in domestic breeds and wild goat species reared in Sudan [15]. Nowadays, molecular mutations, from single nucleotide substitutions/deletions to large insertions/deletions, cause most of the above-mentioned different amounts of αs1-Cn, β-Cn, and αs2-Cn variants in the milk. They are known and quite different, and various research groups have developed quick and cheap tests for genotyping other *loci* [2,3,9,10,13,14,16–18]. From the first identification to today, the genetic quali-quantitative

polymorphism of goat calcium-sensitive caseins has raised considerable research interest because it is related to the quality, yield, and composition of milk rather than to the technological and functional properties [19–27].

Therefore, the aims of the current study are (1) to set up a new efficient and specific PCR-based genotyping protocol to detect the *CSN2*^{01} allele in the goat species, (2) to investigate the distribution pattern of known variants of *CSN1S1*, *CSN2*, and *CSN1S2 loci* in Bianca Monticellana, Capestrina, and Ciociara Grigia, and (3) to investigate their association with the parameters that might affect the milk's traits. Bianca Monticellana, Capestrina, and Ciociara Grigia are three endangered goat breeds reared in the Lazio Region (Central Italy). The population sizes are as follows: for the Bianca Monticellana goat, 1556 total heads, including 50 males and 1506 females; for the Capestrina goat, 739 heads, including 31 males and 708 females, and for Ciociara Grigia, 444 heads, including 24 male and 420 female goats (the data was provided by Assonapa 2022). The aptitude of these breeds is mainly milk production, intended for cheesemaking, such as Marzolina cheese, a traditional hard cylindrical-shaped cheese made exclusively from the milk of goats reared in the wild. It is eaten fresh or aged and, during the last years, has been included among Traditional Agri-food Products (TAPs) of the Lazio Region.

2. Materials and Methods

2.1. Farms and Animals

Five farms located in the Latina and Frosinone Provinces, close to each other, were included in this study. The animals were reared following the same traditional management practices of the area: the goats are left to graze in daylight hours (6/8 h/day) and return to the shed at sunset, leaving the dams with their kids at night. The kids are breastfed for up to 45 ± 5 days postpartum. Manual milking is carried out twice a day, in the morning and in the afternoon, starting from weaning to drying off (at about five months). Information concerning the parity number was also available. A total of 188 goats (125 Monticellana, 27 Capestrina, and 36 Ciociara Grigia) were used for the present study. All the animals were enrolled in the Official Birth Register (ASSONAPA) and minimally related. Blood samples were collected (19 males and 169 females) for genetic characterization. Individual milk samples (50 mL) from 169 goats (112 Bianca Monticellana, 27 Capestrina, and 30 Ciociara Grigia, respectively) at three lactation stages (60, 90, and 120-days postpartum, from May to August 2019) were collected in the morning, to evaluate the effect of casein polymorphisms on the milk yield and quality.

2.2. DNA Extraction

DNA was extracted from the blood by use of a Wizard DNA extraction kit (Promega– Madison, WI, USA), following the manufacturer's instructions.

2.3. Genotyping at the CSN1S1 Locus

Genotyping at the *CSN1S1 locus* was carried out in stages with an initial *Xmn*I PCR- RFLP [17] to check which group of alleles the tested animals were carriers of: A*-derived alleles (A, A2, A3, A′, I, G, M, 02, H, and 01), B* alleles (B′, B1, B2, B3, C, D, D1, L, and E), and F or N alleles. DNA from subjects found to carry the A* group allele was subsequently analyzed by AS-PCR [18] to verify if they carried the 01 allele, while DNA from subjects found to carry the B* group allele was genotyped by the technique described by Jansá Perez et al. [16] to distinguish subjects in which the E allele was present.

2.4. Genotyping at the CSN1S2 Locus

Genotyping of the D and 0 alleles at the *CSN1S2 locus* was performed by the *Nco*I PCR-RFLP method, according to [14].

2.5. Genotyping at the CSN2 Locus

In order to identify carriers of the goats' *CSN2*^{01} allele at the DNA level, a new AS-PCR was set up. Sequence primers used for the AS-PCR are listed in Table S1. All primers were designed with DNASIS-Pro version 3.0 software (Hitachi, Tokyo, Japan) using the goats' *CSN2* sequences as templates (GeneBank, nos. AJ011018, AJ011019.3). The length of the amplified fragment-spanning intron 6 (partial) to exon 7 (partial) is 463/464 bp. Amplifications were performed in a 25-µL volume containing 100 ng of genomic DNA, a 1× PCR buffer, 1.5 mM of MgCl2, 200 µM of each dNTP, 10 pmol of each primer, and 1 U of GoTaq® G2 Flexi DNA Polymerase (Promega–Madison, Fitchburg, WI, USA).

The thermal conditions were: 97 °C for 2 min, 30 cycles at 94 °C for 30 s, annealing at 52.5 °C for 45 s, and extension at 72 °C for 1 min. A final extension was carried out at 72 °C for 10 min.

The amplification products were later verified by electrophoresis on a 2% agarose gel (Bio-Rad, Hercules, CA, USA) in a 0.5X TBE buffer and stained with SYBR®green (Lonza Rockland, Inc., Rockland, ME, USA).

2.6. Milk Analyses

The amount of milk (milk yield) from the morning milking was recorded on the farm. The quality parameters analyzed were: percentages of fat, protein, total casein, lactose and solids-not-fat, and amount of urea and somatic cells. Analyses were performed at the Milk Laboratory of Istituto Zooprofilattico Sperimentale Lazio e Toscana (IZSLT), the Latina section, using COMBIFOSS 6000® (Foss, Hillerød, Denmark) automated equipment, consisting of:

- - Milkoscan FT 6000 (the i.r. spectrophotometry method for: fat, protein, total caseins, lactose, solids-not-fat, and urea, according to the International Dairy Federation (IDF) standard 141:2013 (ISO-IDF, 2013) [28].
- - Fossomatic FT 5000 (the fluoro-optoelectronic method) for somatic cells, according to the IDF 148–2:2006 method (ISO-IDF, 2006) [29].

Moreover, fat, protein, total caseins, lactose, and solid-not-fat yields were calculated.

2.7. Statistical Analysis

Allele frequencies were calculated by simple allele counting [30]. Possible deviations of genotypic frequencies from expectations were tested by a chi-square test to verify if the population was in the Hardy–Weinberg equilibrium. Moreover, some population genetic indices, namely gene heterozygosity (He), gene homozygosity (Ho), effective allele numbers (Ne), and the Fixation Index (FIS), were obtained by POPGENE32 software version 1.32 (PopGene: Microsoft Window-Based Freeware for Population Genetic Analysis, Edmonton, AB, Canada) [31]. The Polymorphic Information Content (PIC) was calculated according to Botstein et al. [32]. A first statistical analysis was carried out to estimate the effect of detected polymorphisms on the milk yield and composition traits of all the animals considered as a single population. In detail, a mixed model for the repeated measures [33] implemented with SAS software (SAS 9.2 Institute, Inc., Cary, NC, USA) was used to assess the possible relationship between *CSN1S1* polymorphisms and performance traits under study. Milk production data were considered as repeated measures, and the correlations between the measures in the same individual were considered in the statistical model. The statistical model included the genotype as the fixed effect (six levels), days in milk as the fixed effect of the lactation stage (3 intervals of 30 days each), the fixed effect of the breed (three levels), the fixed effect of the parity (two levels, 1st-2nd and 3rd and later), the random animal effect, and the residual error term. Subsequently, the same animals were grouped based on their genotype at the *CSN1S1 locus* into three different clusters, called strong (animals carrying A*A*, A*B*, or B*B* genotypes), intermediate (animals carrying FA* or FB* genotypes), and weak (FF goats). The model used was the same model described above. The values were considered significant at $p < 0.05$ and presented as the least squares

means ± standard errors in both cases. If more than two groups were compared, a Bonferroni test was used for multiple testing.

3. Results

3.1. A New AS-PCR for the CSN2^{01} Allele Detection

A new fast and economical method of analysis, based on AS-PCR, was set up to identify carriers of the SNP at position 373 on the seventh exon (AJ011018:g.8915C>T) that characterizes the *CSN2*^{01} allele [10]. For this purpose, two different allele-specific reverse primers (named CSN2N and CSN201) that differ in the last nucleotide at the 3′-end (G→A) (Table S1) were designed. Thus, for the samples without the *CSN2*^{01} allele, PCR amplification was successful only using the reverse primer with guanine at the 3′-end, whereas the *CSN2*^{01} homozygote samples were successfully amplified only by the reverse primer with adenine at the 3′-end. The heterozygote samples were effectively amplified with both reverse primers (Figure 1 and Figure S2). The common forward primer (named CSN2) and the allele-specific reverse primer's sequence are part of intron 6 and exon 7, respectively, and the amplified fragment length is 463/464 bp.

3.2. Genotyping

No animals were found to be carriers of the *CSN1S1*^{01}, *CSN1S1*^{E}, *CSN2*^{01}, *CSN1S2*^{D}, and *CSN1S2*^{0} alleles, applying the methods already known for the *CSN1S1* and *CSN1S2* genes and the new AS-PCR protocol for *CSN2*. Table 1 shows the allele and genotype frequencies at the *CSN1S1 locus*.

Among the six allele classes investigated at the *CSN1S1 locus*, four of them were found in the studied population: A*, B*, F, and N. *CSN1S1*^{F} (0.44) was the most common, and it was followed by alleles from group B* (0.30) and those from group A* (0.26). The *CSN1S1*^{N} allele was present in the heterozygous state in only one subject of the Bianca Monticellana breed. This finding suggests that it is almost absent in Latium goats and, therefore, cannot be considered typical of these breeds. The application of the *CSN1S1* allele discrimination for the population variability evaluation is informative, being PIC, and calculated for the four alleles found: 0.58 in Bianca Monticellana, 0.51 in Capestrina, 0.58 in Ciociara Grigia, and 0.58 in overall the population.

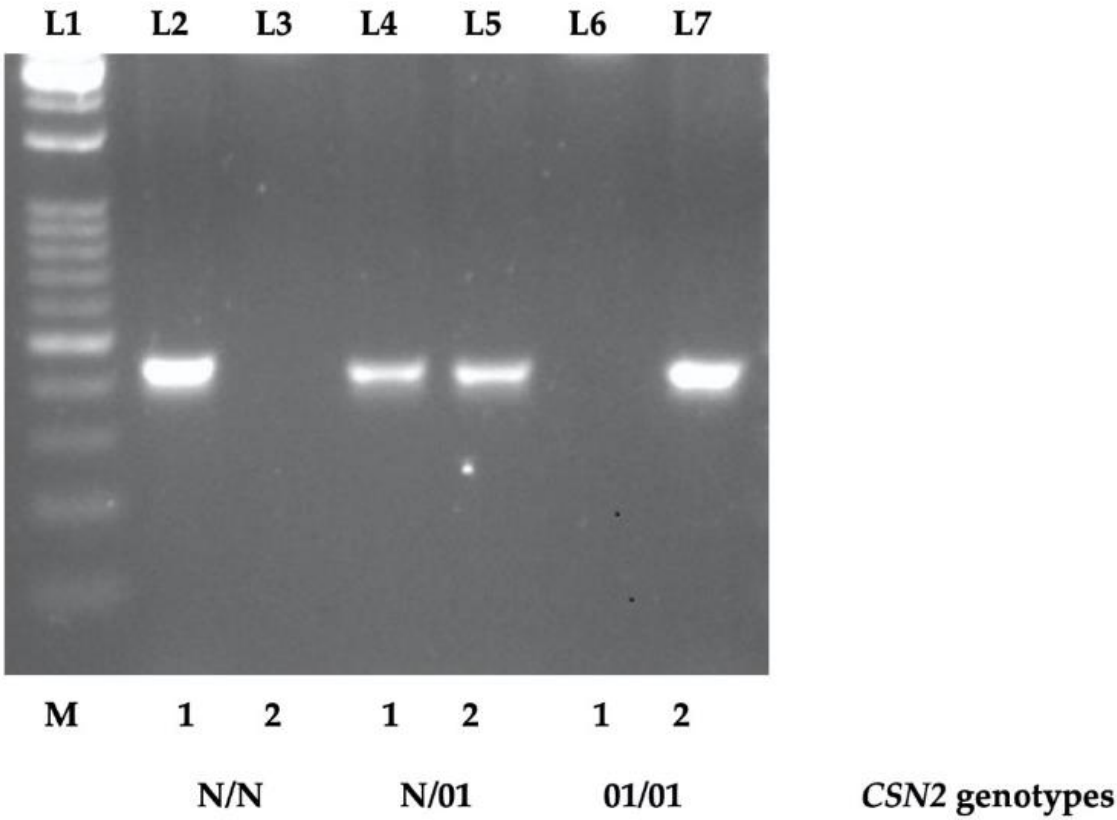

Figure 1. Electrophoretic patterns of fragments obtained with the new AS-PCR protocol of the goats carriers of AJ011018:g.8915C>T mutation at CSN2 locus. M= marker (1kb Opti-DNA Ladder, 0.1–10 kb, Applied Biological Materials, ABM). 1, "g.8915C" allele specific primer (primer name CSN2N), ; 2, "g.8915T" allele specific primer (primer name CSN201) N: non $CSN2^{01}$ alleles (A, A1, B, C, C1, D, E, F and 0), 01: $CSN2^{01}$ allele N/N, N/01 and 01/01: reference samples of individuals with known genotype. Lane 1: Marker; Lanes 2 and 3: N/N goat; Lanes 4 and 5: N/01 goat, Lanes 6 and 7: 01/01 goat.Typing of the reference samples (2 CSN2 N/01 and 2 CSN2 01/01) has been accomplished eight times with identical results.

Table 1. Genotype frequencies (observed and expected), allele frequencies and population indices observed at *CSN1S1 locus* in Capestrina, Bianca Monticellana and Ciociara Grigia populations.

Breed	Number	Genotype frequency							Allele frequency				Population indices		
		A*A*	A*B*	B*B*	FF	B*F	A*F	FN	A*	B*	F	N	Ho	He	F_{IS}
Capestrina n=27	Obs	1	2	2	9	8	5	0	0.170	0.260	0.570	0	0.42	0.58	0.14
	Exp	0.75	2.33	1.81	8.90	8.04	5.17	0							
	χ^2=0.16 p=0.01 d.f.=5														
Bianca Monticellana n=125	Obs	15	18	9	24	35	23	1	0.284	0.284	0.428	0.004	0.61	0.66	0.06
	Exp	10.08	20.16	10.08	22.90	30.39	30.39	0.43							
	χ^2=6.06 p=0.01 d.f.=6														
Ciociara Grigia n=36	Obs	1	6	5	3	12	9	0	0.240	0.390	0.380	0	0.75	0.66	-0.14
	Exp	2.01	6.61	5.44	5.06	10.50	6.38	0							
	χ^2=2.73 p=0.01 d.f.=5														
All breeds n=188	Obs	17	26	16	36	55	37	1	0.26	0.30	0.44	0.003	0.63	0.65	0.03
	Exp	12.51	29.15	16.98	36.20	49.59	42.57	0.44							
	χ^2=4.04 p=0.01 d.f. =6														

A*=A, A2, A3, A′, I, G, M, 02, H, 01; B*=B′, B1, B2, B3, C, D, D1, L, E

3.3. Allele Effect on Milk Parameters

The data reported in Table 2 show the effects of the six different *CSN1S1* genotypes found in the studied population. No differences in terms of milk yield were observed. Significant differences among the genotypes were found in the fat and protein yielded and in the fat, protein, solids-not-fat, and casein contents. In detail, animals carrying the A*A* genotype produced more fat per milking if compared to the A*B* and B*F ones ($p < 0.05$). Similar results were found for the protein yield, with A*A* individuals being more productive than goats with the B*F genotype ($p < 0.05$). Moreover, A*A* and B*B* individuals produced milk with a greater fat content than their A*B* and FF counterparts. As per the protein, casein, and solids-not-fat percentage, the data show a statistically significant difference among the genotypes, with milk produced by the FF goat having a lower content of these constituents if compared with all the other genotypes.

Table 3 shows the effects of different CSN1S1 genotypes clustered according to the αs1-Cn content on milk yield and composition. Significant differences were observed in the fat, protein, solids-not-fat, and casein percentages. The animals carrying strong genotypes produced milk with a greater percentage of fat compared to those in the weak group ($p < 0.01$). Moreover, the percentage of the protein and casein were significantly higher in milk produced by individuals with strong and intermediate genotypes if compared with the weak ones ($p < 0.01$), with the animals belonging to the strong group showing a higher value when compared with the intermediate genotypes ($p < 0.05$). Consequently, the milk produced by the strong and intermediate genotypes was richer in solids-not-fat than that produced by a goat carrying a weak genotype ($p < 0.01$).

Table 2. Effect of different genotypes at *CSN1S1 locus* on milk yield and composition of all the animals studied.

Parameters	Genotypes					
	A*A* *n*=16	A*B* *n*=24	B*B* *n*=15	A*F *n*=34	B*F *n*=44	FF *n*=35
Milk yield (g/milking)	608.13±66.26	521.54±51.98	507.14±70.83	534.46±43.58	470.00±39.51	552.57±44.80
Fat yield (g)	29.54±3.05^{a}	21.23±2.39^{b}	25.10±3.26	22.55±2.01	19.96±1.82^{b}	22.41±2.06
Protein yield (g)	22.28±2.27^{a}	18.83±1.78	18.91±2.42	18.80±1.49	16.46±1.35^{b}	17.42±1.53
Lactose yield (g)	26.68±3.00	22.93±2.35	22.26±3.20	23.35±1.97	20.71±1.79	24.49±2.03
Solids-not-fat (g)	53.08±5.64	45.21±4.42	44.49±6.03	45.85±3.71	40.26±3.36	45.46±3.81
Casein (g)	16.78±1.67	14.10±1.33	14.25±1.81	14.02±1.11	12.30±1.01	12.77±1.14
Fat (%)	4.90±0.26^{a}	4.06±0.20^{b}	4.97±0.26^{a}	4.33±0.17	4.32±0.15	4.08±0.17^{b}
Protein (%)	3.67±0.09^{A}	3.65±0.07^{A}	3.78±0.09^{A}	3.56±0.06^{A}	3.50±0.05^{a}	3.22±0.06^{Bb}
Lactose (%)	4.38±0.06	4.35±0.05	4.40±0.06	4.33±0.04	4.44±0.03	4.43±0.04
Solids not fat (%)	8.73±0.10^{a}	8.66±0.08^{a}	8.85±0.10^{A}	8.60±0.07^{a}	8.61±0.06^{a}	8.31±0.07^{Bb}
Casein (%)	2.76±0.08^{A}	2.72±0.06^{A}	2.86±0.08^{A}	2.66±0.05^{A}	2.62±0.05^{A}	2.38±0.05^{B}
Urea (mg/100ml)	48.94±2.14	49.36±1.68	47.46±2.14	50.05±1.40	51.63±1.25	50.53±1.39
Somatic cell count (x10^{3})	1909.13±689.73	1616.04±541.07	1602.31±689.73	2006.89±466.35	1610.36±402.43	1598.94±447.55

A, B=p<0.01; a, b=p<0.05; A*=A, A2, A3, A′, I, G, M, 02, H, 01; B*=B′, B1, B2, B3, C, D, D1, L, E

Table 3. Effect of genotypes at *CSN1S1 locus* clustered as Strong, Intermediate and Weak on milk yield and composition of all the animals studied.

Parameters	**Cluster**		
	Strong *n*=55	**Intermediate *n*=78**	**Weak *n*=35**
Milk yield (g/milking)	542.68±35.37	499.09±29.23	552.57±44.74
Fat yield (g)	24.57±1.64	21.13±1.36	22.41±2.08
Protein yield (g)	19.83±1.21	17.51±1.00	17.42±1.53
Lactose yield (g)	23.84±1.60	21.90±1.32	24.49±2.02
Solids not fat (g)	47.28±3.01	42.78±2.49	45.46±3.81
Casein (g)	14.90±0.90	13.08±0.75	12.77±1.14
Fat (%)	4.54±0.14^{A}	4.32±0.12	4.08±0.17^{B}
Protein (%)	3.69±0.05^{Aa}	3.53±0.07^{Ab}	3.22±0.09^{B}
Lactose (%)	4.37±0.03	4.39±0.03	4.43±0.04
Solids-not-fat (%)	8.73±0.05^{A}	8.61±0.05^{A}	8.31±0.07^{B}
Casein (%)	2.77±0.04^{Aa}	2.64±0.03^{Ab}	2.38±0.05^{B}
Urea (mg/100 ml)	48.72±1.12	50.94±0.93	50.53±1.38
Somatic cell count (x10^{3})	1693.10±359.72	1779.61±302.53	1598.94±444.41

A, B=p<0.01; a, b=p<0.05

Strong genotypes =A*A*, A*B*, B*B*; Intermediate genotypes = A*F, B*F; Weak genotype= FF.

4. Discussions

4.1. AS-PCR Protocol for CSN201 Allele Detection

From the first goat milk protein polymorphisms described by Boulanger et al. [34] for αs1 and αs2-Cn and by Dall'Olio et al. [35] for β-Cn through the application of the electrophoretic technique, an increasing number of protein variants have been discovered over the years. The development of molecular genetics technologies in recent decades has also made it possible to obtain an extraordinary amount of information about the animal genome enabling the identification of causative events of the observed phenotypic differences and the identification of new alleles.

The later application of genotyping methods (PCR-RFLP, AS-PCR, ACRS-PCR, SSCP, DGGE . . .) has allowed a rapid and economical genotyping of individuals, regardless of the phenotypic expression, sex, and age. Identifying new alleles and mutations requires a review of existing genotyping protocols to optimize them, considering the latest findings.

From this perspective, it was necessary to set up a new AS-PCR reaction to detect carriers of the *CSN201* allele in the goat. In fact, it has been observed that both allele-specific forward primers (5′-CGTGCTGTCCCTTTMTC -3′ and 5′- CGTGCTGTCCCTTTMTT -3′) proposed by Ramunno et al. [36] include, in the third-to-last nucleotide at the 3′-end, the transversion AJ011018.3:g.8913C>A responsible for the amino acid exchange, p.Ser166>Tyr, in the mature protein encoded by the most recently identified allele, *CSN2E* (Figure S2). This condition could reduce the efficiency and specificity of the reaction by making the genotyping method proposed by Ramunno et al. [36] ambiguous in the presence of *CSN2E* variant carriers. Hence, using the method developed in this study, it is now possible to quickly genotype goats at the *CSN2 locus* precisely and unequivocally.

4.2. Milk Traits Phenotyping

The qualitative parameters of the milk of Lazio goats, as regards the percentages of proteins, fat, lactose, not-fat-solids, and somatic cells, are perfectly aligned with those of other native and highly selected Alpine breeds [37,38]. However, considering the differences in terms of the milk yield between native and cosmopolitan breeds, the total productions per milking (the fat, protein, lactose, caseins, and solids-not-fat yields), these goats are certainly more similar to the first ones [39,40]. These findings show that the breeding of Lazio goats has its validity in consideration of the low breeding costs and that they allow the recovery of areas where other types of livestock activities would not be economically sustainable.

4.3. Calcium-Sensitive Caseins Loci Genotyping and Population Genetic Structure

To assess the possible correlations between the *CSN1S2* genotype and the milk parameters in the investigated goat populations, a genotyping was carried out to detect the null (0) and intermediate (D) alleles at this locus. No carriers of both the 0 and D alleles were found. This result was expected, as both of these alleles are rare and detected only in a few goat breeds. In fact, from the discovery of the 0 allele in the Napoletana goat breed [14], it was mainly observed in Italian goat breeds, such as Argentata dell'Etna [41], Maltese, Jonica [42], and Sarda [43], as well as in Saanen reared in the Bursa Province in the Marmara Region of Turkey [44] and in some local Hungarian breeds [45]. Only in these latter populations, exceptionally, the *CSN1S20* allele was observed with a relatively high incidence (0.146). Even rarer is the *CSN1S2D* allele, which is still identified only in Napoletana goats (0.019) [14] and Hungarian breeds (0.005) [45].

Similarly to what was observed at the *CSN1S2 locus*, the analysis of the *CSN2 locus* showed the absence of the null allele, *CSN2*01, in the investigated populations. This result is compatible with the ones obtained in other breeds reared in Italy and characterized by the lack or very low frequencies of this null allele [10,24,36,42,43,46,47].

In this study, no analyses have been performed to detect the null allele, *CSN20*, that, similarly to the 01 allele, is characterized by a premature stop codon (codon 58) due to a single nucleotide deletion (adenine) in a row of four adenines between nt 16 and 19 of exon 7. From its identification and characterization by Persuy et al. [48] in the Pyrenean goat breed, the presence of this allele was, in fact, no longer reported in any other goat breed.

Molecular analyses showed a fair genetic variability at the *CSN1S1 locus* in the goat populations studied. Four alleles (αs1-Cn A*, B*, F, and N) and seven of the sixteen possible genotypes (Table 1) were found. In particular, this study showed the absence in these populations of alleles associated with an intermediate or absent content of this protein in the milk, with only one exception for one Bianca Monticellana goat, where we found a heterozygote for the allele N (*CSN1S1* F/N). Conversely, grouping the three genetic types, a high frequency of alleles associated with a low (*CSN1S1F*) and high (*CSN1S1A*,B**) content of αs1 casein content in the milk, with *CSN1S1F* ≥ *CSN1S1B** ≥ *CSN1S1A**, was observed (Table 1).

We compared the allele frequencies measured in the present work with those reported in different studies on goat populations reared in Italy and genotyped with the same or comparable techniques in this study. This comparison shows that these Lazio goat breeds have an intermediate genotype between the goat breeds reared in Northern Italy (Frisa, Orobica, Verzasca, Vallesana, Saanen, and Roccaverano), which are characterized by a high frequency of alleles associated with a null or low/intermediate αs1-Cn content. Instead, in the authocthonous breeds of Southern Italy, the alleles

*CSN1S1B** and *CSN1S1A** are predominant (Table S2). The Napoletana goat breed is an exception within this panorama, and it stands out from the other breeds for the high frequency of the null allele *CSN1S1N* and the *CSN1S1F* allele (Table S2). This distinct genetic structure could be the consequence of geographic isolation after domestication. The Napoletana goat is mainly reared in the Lattari Mountains (the Campania Region), and it could have better preserved ancient alleles or variants rare or absent in other populations. On the contrary, in some breeds of Alpine origin (Saanen and Alpine Italian Chamois) the genetic selection of recent years is causing an increase in the frequency of strong and intermediate alleles at the expense of those of the weak and null [49].

For the remaining breeds, there are no genetic improvement actions. Consequently, the allelic frequencies remain somewhat stable over time, giving well-defined genetic structures to these breeds, primarily responsible for the particular chemical–physical, technological, and organoleptic characteristics of the milk produced and their derivatives.

4.4. Association Study

The polymorphisms of *CSN1S1* affect not only the quantity of casein in goat milk but also its structural and nutritional characteristics (the diameter of the casein micelles, calcium content, fat, fatty acid profile, and urea level) [50–55], and technological properties (the coagulation parameters, cheese yields, and organoleptic properties) [22,56]. Moreover, a greater digestibility of goat milk containing αs1-Cn weak or null content has been shown [19].

In the present study, we have found an association between different genotypes at the *CSN1S1 locus* and some milk traits, namely the fat and protein yielded and fat, protein, solids-not-fat, and casein percentages, without having an effect on the milk yield. Similar results were observed when the goats were grouped into three clusters (strong, intermediate, and weak genotypes).

This is in line with the literature being similar to those reported by [52], obtained by comparing goats with different *CSN1S1* genotypes. These authors grouped the animals into two separate clusters called high (A, B, and C alleles) and low (F, G, 0, and E alleles). They found that milk produced by the low group has lower protein and fat contents than the high group, with no difference in the milk yield and lactose concentration. Moreover, Balia et al. [51] reported a lack of association between the *CSN1S1* genotype and the milk yield in a flock of Sarda goats. Conversely, the milk protein and casein percentages were significantly affected by the genotype: the milk obtained by BB individuals was characterized by a high protein percentage than that of the AF and BF. High amounts of αs1-Cn expressed by the *CSN1S1* BB genotype have also been reported in Cilentana goats without an effect on the milk protein and total casein concentrations [27].

Interestingly, goat breeds in this report show a trend to higher protein and casein percentages in homozygote B* goats than homozygote A* ones at the *CSN1S1* gene. In support of that, Montalbano et al. [57] refer that quantifying by means of RP-HPLC B* genetic variant compared to A* show that the expression of this allele determines a higher content of αs1-casein in Girgentana goat milk.

A different degree of expression between the *CSN1S1* A* and B* alleles may be a consequence of the presence of rare and not well-characterized A* alleles associated with a lower synthesis level, such as I (intermediate) or G (low) or 02 (null) [2]. The applied genotyping method by means of *Xmn*I PCR-RFLP does not differentiate among these alleles. Another possible hypothesis is that there are individuals with B* allele variants, such as *CSN1S1*B3, with significant effects on the protein and casein percentages [58].

Furthermore, variations in the upstream region of the *CSN1S1* gene may affect the protein expression and significantly affects the protein percentage [59–62]. In particular, a binding site of the Activator Protein (AP-1), known to be the critical third messenger for the target genes, regulated by extracellular mediators, and involved in the gene regulation of the mammary epithelial cells, as a response to prolactin, is affected by an A→G exchange at −175 bp in the bovine *CSN1S1* promoter, associated to variations in the expression of the corresponding gene product [62]. Similarly, Ramunno et al. [7] refer that the mutation AJ504711:g653A>G seems to create an extra AP-1 binding motif in the proximal promoter sequence of the goat's *CSN1S1* B* derived alleles. Therefore, it is possible to hypothesize that the mutation could be responsible for the observed expression level changes of goats' strong alleles (A* e B*). Currently, studies are ongoing to verify this hypothesis (Cosenza, unpublished data).

5. Conclusions

The opportunity to accurately characterize the genetic structure of *loci* of interest in livestock species allows us to better plan the breeding and selection activities. The new genotyping technique reported in this paper enables a more accurate characterization of the *CSN2 locus* in goats. Future aims include the development and application of an analysis protocol for each variant at the casein loci to rate their effect on milk traits.

Moreover, in this study, the autochthonous goat breeds of the Lazio Region have been genetically characterized for the first time at the quantitative alleles of calcium-sensitive caseins. These data are essential for the correct genetic management of these breeds to avoid the modification of their population's genetic structure and to guarantee the preservation of typical productions over time. Moreover, such breeds play two crucial functions, allowing the use of marginal areas, avoiding their abandonment, and preserving genetic variability and biodiversity.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ani13020239/s1, Figure S1: (a) flock of the Bianca Monticellana goat breed that goes to pasture; (b) the Capestrina goat at recovery; (c) the Ciociara Grigia goat breed at the pasture; Figure S2: Goat *CSN2* exon 7 nucleotide sequence (capital letters), plus the parzial intron flanking region (small letters). Primer sequences for AS-PCR reported by Ramunno et al. [36] are double underlined. Amino acid sequence is in uppercase and bold letters. Primer sequences for AS-PCR reported in this work are boxed. The transversion AJ011018.3:g.8913C>T responsible for the amino acid exchange, p.Ser166>Tyr, in the mature protein encoded by the $CSN2^{E}$ allele is highlighted in gray. The $CSN2^{01}$ premature stop codon is symbolized by *; Table S1: Primer sequences and position used for AS-PCR; Table S2: Comparison of the allele frequencies for the *CSN1S1 locus* in different Italian goat breeds. References [63,64] are cited in the supplementary materials.

Table S1. Primer sequences and position used for AS-PCR

***CSN2*-amplified fragment**	**Position**	**Primer sequence (5′-3′)**	**Primer name**
Partial intron 6 - partial exon 7	8470-8495 *	F: AGGAGGATAATTCATCATGAATAACA	CSN2
	Complementary to 8914 - 8932 **	R: CAGGCAGAACTTTGGGCTA	CSN201
	Complementary to 8915- 8932 *	R: AGGCAGAACTTTGGGCTG	CSN2N

* Numbering of primers was according to the goat *CSN2* sequence (AJ011018.3)
** Numbering of primer is according to the goat *CSN2* sequence (AJ011019.3)

Table S2. Comparison of the allele frequencies for the *CSN1S1 locus* in different Italian goat breeds.

Goat breed	***CSN1S1***						**References**
	F	N	B*	A*	E	01	
Capestrina	0.5700	-[1]	0.2600	0.170	-	-	Present work
Bianca Monticellana	0.4260	0.004	0.2800	0.2800	-	-	Present work
Ciociara	0.3800	-	0.3900	0.2400	-	-	Present work
Napoletana	0.3680	0.227	0.1800	0.1420	0.0830	-	[2]
Saanen	0.1400	-	0.0600	0.3500	0.4300	0.0200	[49]
Camosciata Alpina	0.1100	-	0.1500	0.5600	0.1800	-	[49]
Girgentana	0.2900	0.047	0.0650	0.5900	0.008	-	[63]
Sarda	0.2420	0.006	0.5200	0.2020	0.0300	0.0010	[25]
Argentata dell'Etna	0.3690	-	0.6310		-	-	[41]
Rossa Mediterranea	0.480	-	0.5200		-	-	[64]
Vallesana	0.3860	-	0.1270	0.0300	0.2830	0.1750	[42]
Roccaverano	0.3770	0.0190	0.1170	0.2270	0.2140	0.0450	[42]
Jonica	0.2820	-	0.3050	0.3500	0.0640	-	[42]
Garganica	0.2281	0.0088	0.2807	0.4298	0.0175	0.0351	[56]
Maltese	0.3710	-	0.1570	0.4140	0.0570	-	[42]
Frisa	0.5570	-	0.0070	0.1290	0.2000	0.1070	[46]
Orobica	0.9620	-	0.0080	-	0.0080	0.0230	[46]
Verzasca	0.7540	-	0.0370	-	0.2010	0.0070	[46]

[1]-: allele frequency = 0.

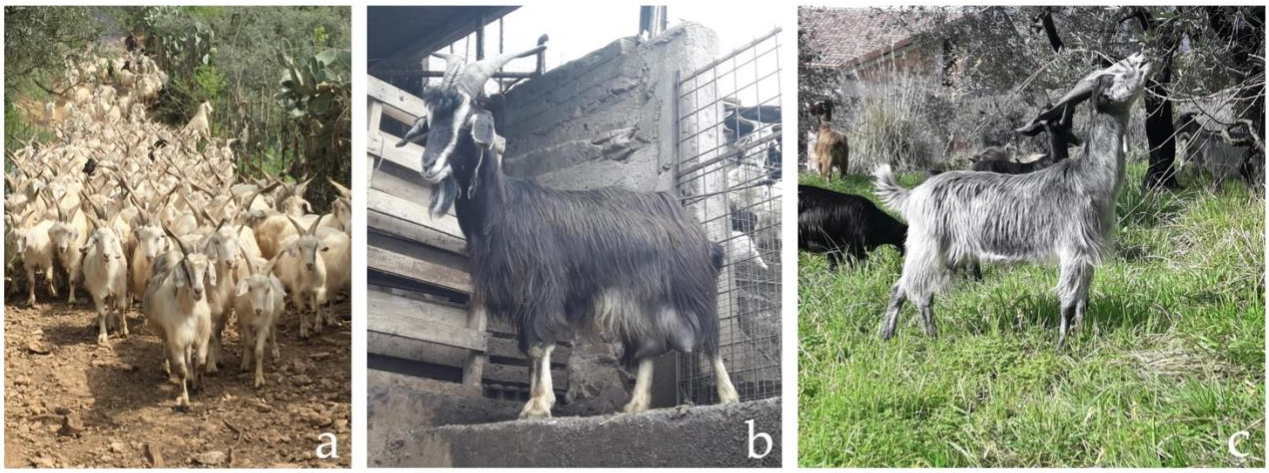

Figure S1: a) flock of Bianca Monticellana goat breed that goes to pasture b) Capestrina goat at the recovery; b) Ciociara Grigia goat breed at the pasture

aggaggataattcatcatgaataacaattataactggattatggactcaaagatttttttccttctttccag

D E L Q D K I H P F A Q A Q S L V Y P F T G P I P N S
GATGAACTCCAGGATAAAATCCACCCCTTTGCCCAGGCACAGTCTCTAGTCTATCCCTTCACTGGGCCCATCCCTAACAG

L P Q N I L P L T Q T P V V V P P F L Q P E I M G V
CCTCCCACAAAACATCCTGCCTCTTACTCAAACCCCTGTGGTGGTGCCGCCTTTCCTTCAGCCTGAAATAATGGGAGTCC

P K V K E T M V P K H K E M P F P K Y P V E P F T E S
CCAAAGTGAAGGAGACTATGGTTCCTAAGCACAAAGAAATGCCCTTCCCTAAATATCCAGTTGAGCCCTTTACTGAAAGC

Q S L T L T D V E K L H L P L P L V Q S W M H Q P P Q
CAGAGCCTGACTCTCACTGATGTTGAAAAGCTGCACCTTCCTCTGCCTCTGGTCCAGTCTTGGATGCACCAGCCTCCCCA

166
P L S P T V M F P P Q S V L S L S Q P K V L P V P Q
GCCTCTTTCTCCAACCGTCATGTTTCCTCCTCAGTCCGTGCTGTCCCTTTCTCAGCCCAAAGTTCTGCCTGTTCCCCAGA

A T
Y *
K A V P Q R D M P I Q A F L L Y Q E P V L G P V R G P
AAGCAGTGCCCCAGAGAGATATGCCCATCCAGGCCTTTCTGCTGTACCAGGAGCCTGTACTTGGTCCTGTCCGGGGACCC

F P I L
TTCCCTATTCTTgtaagtctaaatttactaactgtgctgtttaacttctgatgtttgtatgatatttgagtaattaagag

ccctacaaaaaatcaataatgaatggttccaaaataagcatagctgagattaatgattctcagcattagttataaataga

ataagctggaaaac

Figure S2. The goat *CSN2* exon 7 nucleotide sequence (capital letters) plus the parzial intron flanking region (small letters). Primer sequences for AS-PCR reported by Ramunno *et al*. (1995) are double underlined. Amino acid sequence is in uppercase and bold letters. Primer sequences for AS-PCR reported in this work are boxed. Transversion AJ011018.3:g.8913C>T responsible for the amino acid exchange p.Ser166>Tyr in the mature protein encoded by *CSN2*^{E} allele is highlighted in gray. *CSN2*^{01} premature stop codon is symbolized by *****.

Author Contributions: Conceptualization, F.C. and G.C.; Methodology, S.A., G.C., and M.S.; Formal Analysis, E.D., M.P., M.S., T.G., G.S., and G.C.; Resources, V.P. and T.G.; Funding Acquisition, G.S., T.G., and V.P.; Data Curation, S.A., M.S., A.I., and T.G.; Writing—Original Draft Preparation, G.C., M.S., M.P., and T.G.; Writing—Review and Editing, G.C., M.S., S.A., A.I., and F.C.; Supervision, T.G., F.C., and V.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Italian Ministry of Health with current research funding (RC IZS LT 10.19– coordinator Dr. Remo Rosati, principal investigator Dr. Giorgio Saralli).

Institutional Review Board Statement: This study was carried out following the recommendations of the European Council Directive (86/609/EEC) on the protection of animals. Ethical approval was not required in this study, as the milk and blood samplings were performed by a vet as part of a routine clinical visit.

Informed Consent Statement: Informed consent was obtained from the owners of the animals. **Data Availability Statement:** The data presented in this study are available on request from the

corresponding author.

Acknowledgments: The authors wish to thank M. Brancaleone and C. Digiovannantonio of Agenzia Regionale per lo Sviluppo e l'Innovazione dell'Agricoltura del Lazio (ARSIAL) and the laboratory staff (G. Bruni, E. Crosato, E. Parise, and P. Parisella) of IZSLT "M. Aleandri", the Latina section) for their useful contribution to this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Albarella, S.; Selvaggi, M.; D'Anza, E.; Cosenza, G.; Caira, S.; Scaloni, A.; Fontana, A.; Peretti, V.; Ciotola, F. Influence of the Casein Composite Genotype on Milk Quality and Coagulation Properties in the Endangered Agerolese Cattle Breed. *Animals* **2020**, *10*, 892.
2. Cosenza, G.; Pauciullo, A.; Gallo, D.; Colimoro, L.; D'Avino, A.; Mancusi, A.; Ramunno, L. Genotyping at the CSN1S1 locus by PCR-RFLP and AS-PCR in a Neapolitan Goat Population. *Small Rumin. Res.* **2008**, *74*, 84–90.
3. Caroli, A.; Chiatti, F.; Chessa, S.; Rignanese, D.; Ibeagha-Awemu, E.M.; Erhardt, G. Characterization of the casein gene complex in West African goats and description of a new αS1-casein polymorphism. *J. Dairy Sci.* **2007**, *90*, 2989–2996.
4. Garro, G.; Ferranti, P.; De Pascale, S.; Nicolai, M.A.; Mauriello, R.; Quarto, M.; Pilla, F.; Chianese, L. The occurrence of genetic polymorphism and related non-allelic proteins increases the composition complexity of goat α_{S1}-CN. *Electrophoresis* **2012**, *33*, 2337–2344.
5. Mangia, N.P.; Saliba, L.; Zoumpopoulou, G.; Chessa, S.; Anastasiou, R.; Karayiannis, I.; Sgouras, D.; Tsakalidou, E.; Nudda, A. Goat Milk with Different Alpha-s1 Casein Genotype (CSN1S1) Fermented by Selected Lactobacillus paracasei as Potential Functional Food. *Fermentation* **2019**, *5*, 55.
6. Mestawet, T.A.; Girma, A.; Ådnøy, T.; Devold, T.G.; Vegarud, G.E. Newly identified mutations at the CSN1S1 gene in Ethiopian goats affect casein content and coagulation properties of their milk. *J. Dairy Sci.* **2013**, *96*, 4857–4869.
7. Ramunno, L.; Cosenza, G.; Rando, A.; Pauciullo, A.; Illario, R.; Gallo, D.; Di Berardino, D.; Masina, P. Comparative analysis of gene sequence of goat CSN1S1 F and N alleles and characterization of CSN1S1 transcript variants in mammary gland. *Gene* **2005**, *345*, 289–299.
8. Devold, T.G.; Nordbø, R.; Langsrud, T.; Svenning, C.; Brovold, M.J.; Sørensen, E.S.; Christensen, B.; Ådnøy, T.; Vegarud, G.E. Extreme frequencies of the αs1-casein "null" variant in milk from Norwegian dairy goats—Implications for milk composition, micellar size and renneting properties. *Dairy Sci. Technol.* **2011**, *91*, 39–51.
9. Cosenza, G.; Pauciullo, A.; Gallo, D.; Di Berardino, D.; Ramunno, L. A *Ssp* I PCR-RFLP detecting a silent allele at the goat *CSN2* locus. *J. Dairy Res.* **2005**, *72*, 456–459.
10. Cosenza, G.; Iannaccone, M.; Pico, B.A.; Ramunno, L.; Capparelli, R. The SNP g.1311T>C associated with the absence of *β*-casein in goat milk influences *CSN2* promoter activity. *Anim. Genet.* **2016**, *47*, 615–617.
11. Nicolai, M.A.; Garro, G.; Caira, S.; Mauriello, R.; Quarto, M.; De Pascale, S.; Chianese, L. Occurrence of quantitative genetic polymorphism at the caprine β-CN locus, as determined by a proteomic approach. *Int. Dairy J.* **2021**, *112*, 104855.
12. Erhardt, G.; Jäger, S.; Budelli, E.; Caroli, A. Genetic polymorphism of goat αS2-casein (CSN1S2) and evidence for a further allele. *Milchwissenschaft* **2002**, *57*, 137–140.

13. Lagonigro, R.; Pietrolà, E.; D'Andrea, M.; Veltri, C.; Pilla, F. Molecular genetic characterization of the goat αs2-casein E allele. *Anim. Genet.* **2001**, *32*, 391–393.
14. Ramunno, L.; Cosenza, G.; Pappalardo, M.; Longobardi, E.; Gallo, D.; Pastore, N.; Di Gregorio, P.; Rando, A. Characterization of two new alleles at the goat CSN1S2 locus. *Anim. Genet.* **2001**, *32*, 264–268.
15. Rahmatalla, S.A.; Arends, D.; Ahmed, A.S.; Hassan, L.M.A.; Krebs, S.; Reissmann, M.; Gudrun, A.; Brockmann, G.A. Capture Capture Sequencing to Explore and Map Rare Casein Variants in Goats. *Front. Genet.* **2021**, *12*, 620253.
16. Jansà-Perez, M.; Leroux, C.; Bonastre, A.S.; Martin, P. Occurrence of a Line sequence in the 3′ UTR of the goat α_{S1}-casein E-encoding allele associated with reduced protein synthesis level. *Gene* **1994**, *147*, 179–187.
17. Ramunno, L.; Cosenza, G.; Pappalardo, M.; Pastore, N.; Gallo, D.; Di Gregorio, P.; Masina, P. Identification of the goat CSN1S1F allele by means of PCR-RFLP method. *Anim. Genet.* **2000**, *31*, 342–343.
18. Cosenza, G.; Illario, R.; Rando, A.; Di Gregorio, P.; Masina, P.; Ramunno, L. Molecular characterization of the goat CSN1S101 allele. *J. Dairy Res.* **2003**, *70*, 237–240.
19. Bevilacqua, C.; Martin, P.; Candalh, C.; Fauquant, J.; Piot, M.; Bouvier, F.F.; Manfredi, E.; Pilla, F.; Heyman, M. Allergic sensitisation to milk proteins in guinea-pigs fed cow milk and goat milks of different genotypes. In Proceedings of the VIIth International Conference on Goat, Tours, France, 15–21 May 2000.
20. Marletta, D.; Bordonaro, S.; Guastella, A.M.; Falagiani, P.; Crimi, N.; D'Urso, G. Goat milk with different αS2-casein content: Analysis of allergenic potency by REAST-inhibition assay. *Small Rumin. Res.* **2004**, *52*, 19–24.
21. Martin, P.; Szymanowska, M.; Zwierzchowski, L.; Leroux, C. The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reprod. Nutr. Dev.* **2002**, *42*, 433–459.
22. Selvaggi, M.; Tufarelli, V. Caseins of goat and sheep milk: Analytical and technological aspects. In *Casein: Production, Uses and Health Effects*; Nova Science Publisher: Hauppage, NY, USA, 2012; pp. 1–26.
23. Selvaggi, M.; Laudadio, V.; Cataldo, D.; Tufarelli, V. Major proteins in goat milk: An updated overview on genetic variability. *Mol. Biol. Rep.* **2014**, *41*, 1035–1048.
24. Tortorici, L.; Di Gerlando, R.; Mastrangelo, S.; Sardina, M.T.; Portolano, B. Genetic Characterisation of *CSN2* Gene in *Girgentana* Goat Breed. *Ital. J. Anim. Sci.* **2014**, *13*, 3414.
25. Vacca, G.M.; Dettori, M.L.; Piras, G.; Manca, F.; Paschino, P.; Pazzola, M. Goat casein genotypes are associated with milk production traits in the Sarda breed. *Anim. Genet.* **2014**, *45*, 723–731.
26. Yue, X.P.; Zhang, X.M.; Wang, W.; Ma, R.N.; Deng, C.J.; Lan, X.Y.; Chen, H.; Li, F.; Xu, X.R.; Ma, Y.; et al. The CSN1S1 N and F alleles identified by PCR–SSCP and their associations with milk yield and composition in Chinese dairy goats. *Mol. Biol. Rep.* **2011**, *38*, 2821–2825.
27. Zullo, A.; Barone, C.M.A.; Chianese, L.; Colatruglio, P.; Occidente, M.; Matassino, D. Protein polymorphisms and coagulation properties of Cilentana goat milk. *Small Rumin. Res.* **2005**, *58*, 223–230.
28. *ISO 9622:2013 | IDF 141:2013*; Milk and Liquid Milk Products. Determination of Fat, Protein, Casein, Lactose and pH Content. ISO—International Organization for Standardization: Geneva, Switzerland; IDF—International Dairy Federation: Brussels, Belgium, 2013.
29. *ISO 13366-2:2006 | IDF 148-2:2006*; Milk. Enumeration of Somatic Cells. Part 2: Guidance on the Operation of Fluoro-Opto- Electronic Counters. ISO—International Organization for Standardization: Geneva, Switzerland; IDF—International Dairy Federation: Brussels, Belgium, 2006.
30. Falconer, D.S.; Mackay, T.F.C. *Introduction to Quantitative Genetics*, 4th ed.; Longman Group Ltd.: Essex, UK, 1996.
31. Yeh, F.C.; Yang, R.; Boyle, T.J.; Ye, Z.; Xiyan, J.M. *PopGene32, Microsoft Windows-based freeware for population genetic analysis*; Version 1.32; University of Alberta: Edmonton, Canada, 2000.
32. Botstein, D.; White, R.L.; Skalnick, M.H.; Davies, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* **1980**, *32*, 314–331.
33. Littell, R.C.; Henry, P.R.; Ammerman, C.B. Statistical analysis of repeated measures data using SAS procedures. *J. Anim. Sci.* **1998**, *76*, 1216–1231.
34. Boulanger, A.; Grosclaude, F.; Mahé, M.F. Polymorphisme des caséines α(s1) et α(s2) de la chèvre (*Capra hircus*). *Genet. Sel.* **1984**, *16*, 157–176.
35. Dall'Olio, S.; Davoli, R.; Russo, V. Una nuova variante di caseina caprina. *Sci. Tecnol. Latt.-Casearia* **1989**, *40*, 24–28.
36. Ramunno, L.; Mariani, P.; Pappalardo, M.; Rando, A.; Capuano, M.; Di Gregorio, P.; Cosenza, G. Un gene ad effetto maggiore sul contenuto di caseina β nel latte di capra. In Proceedings of the Atti XI Congresso Nazionale ASPA, Grado (GO), Palermo, Italy, 19–22 June 1995.
37. Currò, S.; Manuelian, C.L.; De Marchi, M.; Goi, A.; Claps, S.; Esposito, L.; Neglia, G. Italian local goat breeds have better milk coagulation properties than cosmopolitan breed. *Ital. J. Anim. Sci.* **2020**, *19*, 593–601.
38. Ferro, M.M.; Tedeschi, L.O.; Atzori, A.S. The comparison of the lactation and milk yield and composition of selected breeds of sheep and goats. *Transl. Anim. Sci.* **2017**, *1*, 498–506.
39. Currò, S.; Manuelian, C.L.; De Marchi, M.; De Palo, P.; Claps, S.; Maggiolino, A.; Campanile, G.; Rufrano, D.; Fontana, A.; Pedota, G.; et al. Autochthonous dairy goat breeds showed better milk quality than Saanen under the same environmental conditions. *Arch. Anim. Breed.* **2019**, *62*, 83–89.

40. Todaro, M.; Scatassa, M.L.; Giaccone, P. Multivariate factor analysis of Girgentana goat milk composition. *Ital. J. Anim. Sci.* **2005**, *4*, 403–410.
41. Marletta, D.; Bordonaro, S.; Guastella, A.M.; Criscione, A.; D'Urso, G. Genetic polymorphism of the calcium sensitive caseins in Sicilian Girgentana and Argentata dell'Etna goat breeds. *Small Rumin. Res.* **2005**, *57*, 133–139.
42. Sacchi, P.; Chessa, S.; Budelli, E.; Bolla, P.; Ceriotti, G.; Soglia, D.; Rasero, R.; Cauvin, E.; Caroli, A. Casein haplotype structure in five Italian goat breeds. *J. Dairy Sci.* **2005**, *88*, 1561–1568.
43. Pazzola, M.; Dettori, M.L.; Pira, E.; Noce, A.; Paschino, P.; Vacca, G.M. Effect of polymorphisms at the casein gene cluster on milk renneting properties of the Sarda goat. *Small Rumin. Res.* **2014**, *117*, 124–130.
44. Deniz, D.; Sena, A.; Hale, S.; Buse, V.; Faruk, B. Identification of the frequency of CSN1S2 gene alleles and the effects of these alleles and parity on milk yield and composition in Saanen goats. *Large Anim. Revi.* **2021**, *27*, 91–96.
45. Kusza, S.; Veress, G.; Kukovics, S.; Jávor, A.; Sanchez, A.; Angiolillo, A.; Bösze, Z. Genetic polymorphism of αs1- and αs2- caseins in Hungarian milking goats. *Small Rumin. Res.* **2007**, *68*, 329–332.
46. Caroli, A.; Chiatti, F.; Chessa, S.; Rignanese, D.; Bolla, P.; Pagnacco, G. Focusing on the goat casein complex. *J. Dairy Sci.* **2006**, *89*, 3178–3187.
47. Chessa, S.; Budelli, E.; Chiatti, F.; Cito, A.M.; Bolla, P.; Caroli, A. Short Communication: Predominance of β-Casein (CSN2) C Allele in Goat Breeds Reared in Italy. *J. Dairy Sci.* **2005**, *88*, 1878–1881.
48. Persuy, M.A.; Printz, C.; Medrano, J.F.; Mercier, J.C. A single nucleotide deletion resulting in a premature stop codon is associated with marked reduction of transcripts from a goat β-casein null allele. *Anim. Genet.* **1999**, *30*, 444–451.
49. Frattini, S.; Nicoloso, L.; Coizet, B.; Chessa, S.; Rapetti, L.; Pagnacco, G.; Crepaldi, P. Short communication: The unusual genetic trend of αS1-casein in Alpine and Saanen breeds. *J. Dairy Sci.* **2014**, *97*, 7975–7979.
50. Avondo, M.; Pagano, R.I.; Guastella, A.M.; Criscione, A.; Di Gloria, M.; Valenti, B.; Piccione, G.; Pennisi, P. Diet selection and milk production and composition in Girgentana goats with different αs1-casein genotype. *J. Dairy Res.* **2009**, *76*, 202–209.
51. Balia, F.; Pazzola, M.; Dettori, M.L.; Mura, M.C.; Luridiana, S.; Carcangiu, V.; Piras, G.; Vacca, G.M. Effect of CSN1S1 gene polymorphism and stage of lactation on milk yield and composition of extensively reared goats. *J. Dairy Res.* **2013**, *80*, 129–137.
52. Chilliard, Y.; Rouel, J.; Leroux, C. Goats alpha-s1-casein genotype influences its milk fatty acid composition and delta-9 desaturation ratios. *Anim. Feed Sci. Technol.* **2006**, *131*, 474–487.
53. Pagano, R.I.; Pennisi, P.; Valenti, B.; Lanza, M.; Di Trana, A.; Di Gregorio, P.; De Angelis, A.; Avondo, M. Effect of CSN1S1 genotype and its interaction with diet energy level on milk production and quality in Girgentana goats fed ad libitum. *J. Dairy Res.* **2010**, *77*, 245–251.
54. Schmidely, P.; Meschy, F.; Tessier, J.; Sauvant, D. Lactation response and nitrogen, calcium, and phosphorus utilization of dairy goats differing by the genotype for αs1-casein in milk, and feed diets varying in crude protein concentration. *J. Dairy Sci.* **2002**, *85*, 2299–2307.
55. Carillier-Jacquin, C.; Larroque, H.; Robert-Granié, C. Including αs1 casein gene information in genomic evaluations of French dairy goats. *Genet. Sel. Evol.* **2016**, *48*, 54.
56. Santillo, A.; Ciliberti, M.G.; D'Angelo, F.; Albenzio, M. The Effect of Alpha s1 Genotype on Some Physiological and Chemical Milk Characteristics in Garganica Goat. *Anim. Front.* **2022**, *3*, 897172.
57. Montalbano, M.; Tortorici, L.; Mastrangelo, S.; Sardina, M.T.; Portolano, B. Development and validation of RP-HPLC method for the quantitative estimation of αs1-genetic variants in goat milk. *Food Chem.* **2014**, *156*, 165–169.
58. Turhan Dinçel, D.; Ardicli, S.; Samli, H.; Balci, F. Determining the frequencies of B1, B2, B3 and E alleles of the CSN1S1 gene and their effects on milk yield and composition in Saanen goats. *S. Afr. J. Anim. Sci.* **2016**, *46*, 180–190.
59. Cieslak, J.; Wodas, L.; Borowska, A.; Pawlak, P.; Czyzak-Runowska, G.; Wojtowski, J.; Puppel, K.; Kuczynska, B.; Mackowski, M. 5′-flanking variants of equine casein genes (CSN1S1, CSN1S2, CSN2, CSN3) and their relationship with gene expression and milk composition. *J. Appl. Genet.* **2019**, *60*, 71–78.
60. Prinzenberg, E.M.; Weimann, C.; Brandt, H.; Bennewitz, J.; Kalm, E.; Schwerin, M.; Erhardt, G. Polymorphism of the bovine CSN1S1 promoter: Linkage mapping, intragenic haplotypes, and effects on milk production traits. *J. Dairy Sci.* **2003**, *86*, 2696–2705.
61. Sanders, K.; Bennewitz, J.; Reinsch, N.; Thaller, G.; Prinzenberg, E.M.; Kühn, C.; Kalm, E. Characterization of the DGAT1 mutations and the CSN1S1 promoter in the German Angeln dairy cattle population. *J. Dairy Sci.* **2006**, *89*, 3164–3174.
62. Kuss, A.W.; Gogol, J.; Bartenschlager, H.; Geldermann, H. Polymorphic AP-1 binding site in bovine CSN1S1 shows quantitative differences in protein binding associated with milk protein expression. *J. Dairy Sci.* **2005**, *88*, 2246–2252.
63. Mastrangelo, S.; Sardina, M.T.; Tolone, M.; Portolano, B. Genetic polymorphism at the CSN1S1 gene in Girgentana dairy goat breed. *Anim. Prod. Sci.* **2013**, *53*, 403–406.

64. Guastella, A.M.; Criscione, A.; Zuccaro, A.; Tidona, F.; Marietta, D.; Bordonaro, S. Genetic polymorphism of CSN1S1 and CSN2 loci in Rossa Mediterranea goat population. *Ital. J. Anim. Sci.* **2009**, *8*, 101–103.

CHAPTER 2

Set up of new PCT-RFLP protocols to detect *DGAT1* g.5553C>T and g.8539C>T SNPs, genotyping and milk traits association studies in Bagnolese sheep

1. Introduction

The genetic diversity of both livestock and agriculture in the Campania Region finds its maximum expression in its agri-food products. In fact, Campania Region boasts, in addition to the better known PDOs (Caciocavallo Silano, Colatura di Alici di Cetara, Mozzarella di bufala, Provolone del Monaco etc.), 593 PAT testifying to a rich cultural heritage that the territory offers and that we have a duty to preserve.

The breeding of Bagnolese sheep (Fig. 1), a dual-purpose (dairy and meat) breed (Peretti, et al., 2013) whose milk is used to produce Pecorino Bagnolese (Fig. 2), one of the 593 PAT products of the Campania Region, is part of this panorama. To date, the animals listed in the Official Birth Register (National Livestock Association - ASSONAPA) comprise 9.584 adults, including 9.194 females and 390 males, distributed in 119 farms located mainly in Avellino and Salerno provinces and, marginally, in Benevento province.

Fig. 1. Bagnolese sheep

Fig. 2. Pecorino Bagnolese cheese.

As is the case for other native breeds, the survival and spread of the Bagnolese sheep is threatened by the homogenization of farming practices and increasing pressure from globalization. Safeguarding

native breeds from both a biological and a cultural and economic point of view is a challenge for animal husbandry, also in view of the environmental problems the sector is facing.

Precisely because of their nature, these breeds, which have been selected within the specific conditions of the geographical regions and human communities to which they have adapted, contribute to the sustainable production of agro-livestock resources. A conscious and targeted safeguard requires knowledge of the production and genetic peculiarities of the breed one wishes to protect. At present, genetic information concerning the Bagnolese sheep is extremely limited, especially as regards genes affecting production. Consequently, the aim of this study is to increase knowledge in this regard, enabling experts in the sector to carry out targeted selection and better management of the animals currently present.

Among the enzymes known to influence the lipid metabolism at mammary gland level, the Acyl-CoA: diacylglycerol–acyltransferase 1 (*DGAT*1) plays an important role because it catalyzes the final committed step in the formation of triglycerides using diacylglycerol (DGA) and acyl-CoA as substrates (Cases *et al*., 1998).

*DGAT*1 is widely expressed in many tissues, with the highest expression levels in the adipose tissue where it controls the triglycerides synthesis, the adipocyte size and the adipose mass. Over-expression of *DGAT*1 is in fact correlated with the increase in the degree of adiposity (Chen *et al*. 2002a), such as a down-expression brings to thinness and resistance to diet-induced obesity (Smith *et al*., 2000). Regarding the degree of fat unsaturation, the lack of *DGAT*1 expression modifies the fatty acid composition in adipose tissue and skeletal muscle, increasing saturated fatty acids (C16:0 and C18:0) and decreasing monounsaturated fatty acids (C16:1 and C18:1) (Chen *et al*., 2002b).

In cattle, *DGAT*1 became a strong functional candidate for milk fat percentage after Smith *et al*. (2000) described that lactation is absent in knockout mice lacking both copies of *DGAT*1. Furthermore, *DGAT*1 is also a positional candidate, because the gene is located within a region on bovine chromosome 14 that contains a QTL for this trait (Coppieters *et al*., 1998; Zhang *et al*., 1998). Currently, the complete sequences of *DGAT*1 gene are available in different livestock species such as cattle (Winter *et al*., 2002, GenBank accession number AJ318490), pig (Nonneman & Rohrer, 2002, GenBank accession number AY116586), river buffalo (Yuan *et al*., 2007, GenBank accession number AY999090), goat (GenBank accession number LT221856.1) and sheep (Scatà *et al*., 2009; GenBank accession number EU178818.1).

In cattle, the *DGAT*1 gene spans a DNA fragment of 8303 bp, including 1470 bp of exonic regions and 6833 bp of intronic regions. Regarding the gene organization and the lenght of introns, the *DGAT*1 gene structure of the other ruminants is similar (Khan *et al*., 2021). In cattle, the identification of a non-conservative AA → GC exchanges in exon 8, responsible of the aminoacid substitution

p.K232A and of an increase of the amount of alternatively spliced *DGAT*1 mRNA, was associated with milk fat percentage, milk yield (Winter *et al*. 2002; Thaller *et al*., 2003; Grisart *et al*. 2004; Kühn *et al*. 2004) and milk fatty acid profile (Schennink *et al*., 2008; Mach *et al*., 2012). In addition, a polymorphism in the 5' regulatory region of this gene, variable number of tandem repeat (VNTR), also showed a strong association with milk fat percentage (Kuhn *et al*., 2004; Furbass *et al*., 2006; Kuehn *et al*., 2007).

Interesting polymorphisms were described also in buffalo, goat ad sheep *DGAT*1 gene and several studies have attempted to associate SNPs at this *locus* with different milk production traits (for review see Khan *et al*., 2021).

In particular, in sheep it was observed an association between the SNP EU178818:g.5553C>T located in intron 2 with milk fat content showed in Sarda, Altamurana and Gentile di Puglia Italian breeds (Scatà *et al*., 2009), while association studies showed a non-synonymous mutation in exon 17 (EU178818:g.8539C>T; p.Ala487) had a significant effect on milk lactose content, fatty acids content of C4:0, C16:1 c9, and the n-6:n-3 ratio in Assaf ewes. In details, animals carrying the CC genotype had greater lactose, C4:0 and C16:1 c9 contents and lower ratio of n-6:n-3 compare to the CT ones (Dervishi *et al*., 2015).

In addition to milk qualitative traits, the same polymorphism has been widely investigated to identify associations with meat production traits of sheep such as carcass, intramuscular fat content, muscle marbling, fat-tail weight and back fat Thickness, meat tenderness (Xu *et al*., 2009; Mohammadi *et al*., 2013; Altwaty *et al*., 2020; Dai *et al*., 2022; Amri *et al*., 2023) or live weights up to weaning age in lambs (Bayram *et al*., 2019).

It has been widely demonstrated that high milk fat and protein yield % are positively correlated to cheese production. Since *DGAT1* is considered a candidate for the genetic variation of milk characteristics in ruminants, in this study two SNPs, already detected in sheep, have been studied in Bagnolese sheep. To this purpose, steps were taken (1) to set up two new genotyping methods based on PCR-RFLP to identify the sheep carriers of *DGAT1* g.5553C>T and g.8539C>T SNPs, (2) to investigate the distribution pattern of the variants for both SNPs of *DGAT1* in Bagnolose sheep and (3) to check their association with some milk parameters.

2. Materials and methods

2.1. Ethical statement

The Ethical Animal Care and Use Committee of University of Naples Federico II pre-approved all procedures used in this research study (Prot. Nr. PG/2022/0146433).

2.2. Farms and Animals

Eleven farms located in the Avellino, Benevento and Salerno provinces were included in this study. The animals were reared following the same traditional management practices of the area: the sheep are left to graze in daylight hours (6/8 h/day) and return to the shed at sunset. The lambs are breastfed for up to 30 ± 5 days postpartum.

Mechanical milking is carried out twice a day, in the morning and in the afternoon, starting from weaning to drying off (at about six months). Information concerning the parity number was also available.

A total of 252 Bagnolese sheep were used for the present study. All the animals were enrolled in the Official Birth Register (ASSONAPA) and minimally related.

Blood samples were collected (19 males and 233 females) for genetic characterization. Individual milk samples (100 mL) from 90 sheep from the third lactation, reared in a pilot farm joining the SAVEPEB enhancement project, each 15 days for five months (from January to June 2022) were collected in the morning, to evaluate the effect *DGAT1* polymorphisms on the milk parameters.

2.3. DNA Extraction

DNA was extracted from the blood by use of a Wizard DNA extraction kit (Promega– Madison, WI, USA), following the manufacturer's instructions.

2.4. PCR-RFLPs for sheep DGAT1 SNPs genotyping

To identify the sheep carriers of *DGAT* 1 g.5553C>T and g.8539C>T SNPs, two genotyping methods based on PCR-RFLP were developed.

2.5. Genotyping at the sheep DGAT 1 g.5553C>T locus

A 767 bp DNA fragment spanning part of the 2th intron and partial exon 3 region of the sheep *DGAT*1 gene was amplified by means of PCR carried out by using iCycler (BioRad, CA, USA) with the following primers: DGAT-2F, forward: 5'- TGCATTTCTGAGCCTGTCATC -3' (nucleotides 5342–5362); DGAT-3R, reverse: 5'- AACCGTGCGTTGCTTAAGATC -3' (complementary to nucleotides 6088– 6108).

The 25-µl PCR reaction mix included: 100 ng of genomic DNA, 50 mM KCl, 10 mMTris-HCl (pH 9.0), 0.1% Triton X-100, 3 mM MgCl2, 200 nmol of each primer, dNTPs each at 400 µM, 0.5 U of Taq DNA Polymerase (Promega, Madison, WI), and 0.04% BSA.

The amplification programs consisted of 31 cycles. The first one was characterized by a denaturation at 97 °C for 2 min, annealing at 63 °C for 45 s and an extension step at 72 °C for 2 min. The next 30

cycles involved a denaturation step at 94 °C for 45 s, annealing at 63 °C for 45 s and extension at 72 °C for 2 min with the exception that in the last cycle the extension time was 10 min.
Digestion of 17 ml of each PCR amplification was accomplished with 10 U of *BamH*I endonuclease (Promega, Madison, WI) for 5 h at 37 °C following the supplier's directions for buffer conditions.

2.6. Genotyping at the sheep DGAT1 g.8539C>T locus

For genotyping the SNP g.8539C>T a method based on the *Msp*I endonuclease was devised.
A 365 bp DNA fragment spanning part of the 16th exon to partial exon 17 region of the sheep *DGAT*1 gene was amplified with the following primers: DGAT-16F, forward: 5'-GCATGATGGCACAGGTGA -3' (nucleotides 8305–8322); DGAT-17R, reverse: 5'-GGAGGCAGCTTTCACCAG -3' (complementary to nucleotides 8652–8669).
PCR reaction mix and thermal conditions were performed as reported above. Digestion of 17 ml of each PCR amplification was accomplished with 10 U of *Msp*I endonuclease (Promega, Madison, WI) for 5 h at 37 °C following the supplier's directions for buffer conditions.
All primers were designed with DNASIS-Pro version 3.0 software (Hitachi, Tokyo, Japan) using the sheep *DGAT*1 sequences as templates (GenBank Acc. No. EU178818.1). All PCR and digestion products were analyzed directly by electrophoresis in 2% TBE agarose gel (Bio-Rad, CA, USA) in 0.5X TBE buffer and stained with SYBR green nucleic acid stain (Lonza Rockland, Inc., USA).

2.7. *Milk Analyses*

The amount of milk (milk yield) from the morning milking was recorded on the farm. The quality parameters analyzed were protein and fat content.
The protein content (% w/v) was determined using the Kjeldahl method (AOAC International. Official Methods of Analysis, 18th ed.; Horwitz, W., Latimer, G., Eds.; AOAC: Gaithersburg, MD, USA, 2005; ISBN 0935584773), while the fat content (% w/v) was measured using the Gerber method (AOAC International. Official Methods of Analysis, 18th ed.; Horwitz, W., Latimer, G., Eds.; AOAC: Gaithersburg, MD, USA, 2005; ISBN 0935584773).

2.8. Statistical Analysis

Population genetics and statistical analyses were carried out on the total number of genotyped animals and on the farm animals under research. Allele frequencies and genetic indices of the population analyzed such allele frequencies and genetic indices of the population analyzed such as observed (Ho) and expected (He) gene heterozygosity, effective number of alleles (Ne) and fixation index (FIS), were obtained with POPGENE32 software version 1.32 (PopGene: Microsoft Window-Based

Freeware for Population Genetic Analysis, Edmonton, AB, Canada) (Yeh *et al.,* 2000). Polymorphic information content (PIC) was calculated with Cervus software.

Statistical analysis was conducted to estimate the effect of the detected polymorphisms on milk production and milk composition of the analyzed animals considered as a single population.

A mixed repeated-measures model (Littell *et al.,* 1998) was used with IBM SPSS Statistics software Version 29.0.1.0 to assess the possible relationship between *DGAT1* polymorphisms and the qualitative milk traits under study.

The animals were grouped according to their genotype at the *DGAT1 locus*. Milk production data were considered as repeated measures. The statistical model includes the genotype as a fixed effect (three levels), the fixed effect of parts (two levels, 1st-2nd and 3rd), the random effect of the animal and the residual error term. Values were considered significant at $p < 0.01$. If more than two groups were compared, Bonferroni's multiple testing was used.

3. Results and discussion

3.1. Two new PCR-RFLPs for sheep DGAT1 SNPs genotyping

Two new fast and economical method of analysis, based on PCR-RFLP, were set up to identify carriers of the SNPs at position 1415 of intron 2 (EU178818.1: g.5553C>T) and on the 147th nucleotide of the seventeenth exon (EU178818.1: g.8539C>T), respectively.

The first transition changes a *BamH*I endonuclease restriction site (G/GATC**C,** bold and underlined the polymorphic site) and would allow the identification of T or C-carriers. Therefore, by means of *BamH*I digestion of PCR products, including part of the 2th intron and partial exon 3 (767 bp) of the sheep *DGAT* 1, homozygous individuals for g.5553T show a single undigested fragment of 767 bp, whereas homozygous individuals for g.5553T have two fragments of 207 and 560 bp. Heterozygous individuals produced a pattern characterized by all 3 restriction fragments (Fig. 1 and Fig. 1S).

Likewise, the transition g.8539C>T removes a *Msp*I endonuclease restriction site (C/**C**GG, bold and underlined the polymorphic site). *Msp*I digestion of a PCR product of 365 bp spanning exon 16 (partial) and 17 (partial), would allow carriers for the presence of citosine to be identified. As a consequence, the PCR product, uncut in the presence of timidine, is now restricted to two fragments of 131 and 234 bp (Fig. 1 and Fig. 2S).

5342 tgcatttctgagcctgtcatctccttgggaagccttccgggtaggggactccatccatgt 5401

5402 cgcctagagaagcatcacttttccacagagccttctgcaacccccgtggggcctgaacct 5461

5462 tgagggtggaggtggtggcccctgcccctgcggagggcagccaggcatctggccccaggc 5521

5522 cactggcaagagctcgttgtgatggaggatcygtcctttgctgctgctgtaggagcggcc 5581

5582 gaggcgggtgggggtgtgagtaggggtggagacccaggcccagcttccccagccctcagg 5641

5642 acaggcccgctctttcccaccaccccaccaagggcgtgggcacaccccgcctctggggat 5701

5702 tgggccccggttggtaagggcggaagcccttgggccggtggcagcgtgcaggcaggcttg 5761

5762 gacttcactggggcttggggctgtcgctgtggccaggggcactgacccgctcagtgggac 5821

5822 ggaggatggctgctgggcagcgggtttcttctgccgtggcggcacaggcacctggggttg 5881

5882 cagttggctccagacgggtgggggctgcgtgcccctgcgcaggcacacaggccataggtg 5941

5942 gggagtctcagagcttggcgtgaggtcccgcagggctgggcctgcaggatggaggctgct 6001

6002 gtcctgagctgtgggtgctggcaggagctggggtgggtgttctggggccgcggctgacag 6061

Exon 3

I L S N A R

6062 cattgtgtccctctctctctattgcag**ATCTTAAGCAACGCACGGTT** 6108

Fig 1S. DNA segment comprising the region of the ovine *DGAT*1 gene between the 2nd intron and exon 3 (partial). The C>T transition occurring at nucleotide 1415 of intron 2 (EU178818.1: g.5553C>T) is highlighted in red. The *BamH*I endonuclease restriction site (G/GATCC) is underlined and boxed. The forward (DGAT-2F) and reverse (DGAT-3R) primers are shaded in gray.

Exon 16

8305 **GCATGATGGCACAG**gtgagcagccctggacccccacctgcgagcccaccccgtgggcgca 8364

Exon 17

I P L A W I V G R F F

8365 gaggctcactcccgtcccatgtccccag**ATCCCGCTGGCCTGGATAGTCGGCCGCTTCTT** 8424

```
          R  G  N  Y  G  N  A  A  V  W  L  S  L  I  I  G  Q  P  V  A
8425  CCGTGGCAACTATGGCAACGCGGCTGTGTGGCTGTCACTCATCATTGGGCAGCCAGTGGC  8484

          V  L  M  Y  V  H  D  Y  Y  V  L  N  R  E  A  P  T  A487G  T
8485  CGTCCTGATGTACGTCCACGACTACTACGTGCTCAACCGCGAGGCCCCAACAGCYGGCAC  8544

          *
8545  CTGAGCCCCTCCAGGCTGGTTCCCTCAGGGGTGTTGGACTCCTTTGCCTCACCACTTGCT  8604

8605  GCTGTACTGGAGCCTGCCCCCAACCTGGGCGTAGGGGAGGGGCCTGGCTGGTGAAAGCTG  8664

8665  CCTCC  8669
```

Fig 2S. DNA segment comprising the 3' end of exon 16, intron 16, exon 17, and 3' UTR of the ovine DGAT1 gene. The C>T transition occurring at the 147th nucleotide of exon 17 (EU178818.1: g.8539C>T) is highlighted in red. The *Msp*I endonuclease restriction site (C/CGG) is underlined and boxed. The forward (DGAT-16F) and reverse (DGAT-17R) primers are shaded in gray.

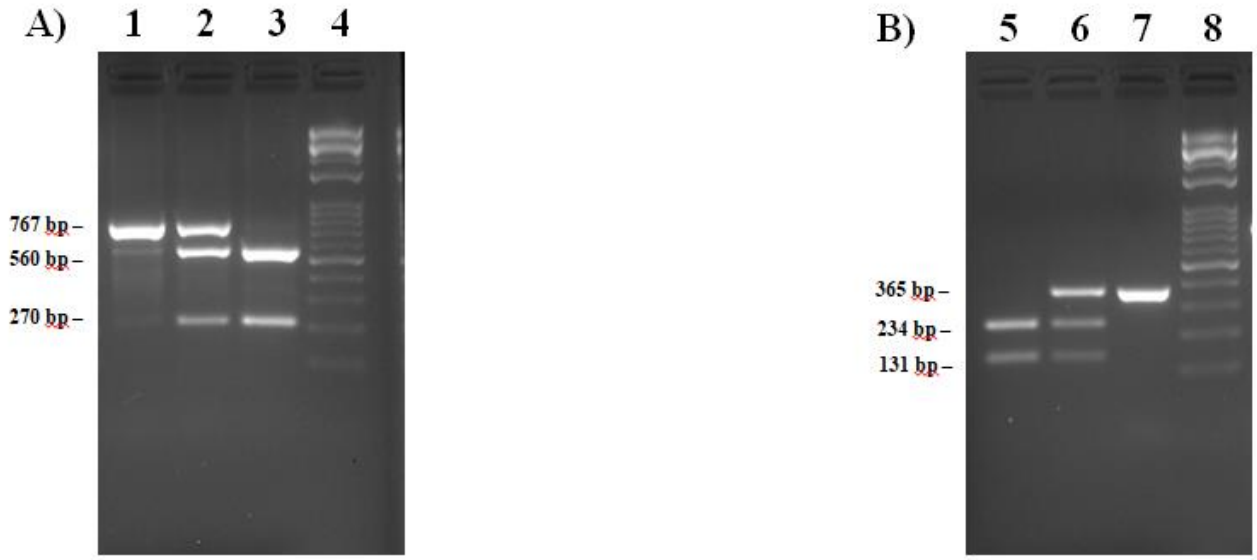

Figure 1. **(A)** Genotyping of the SNP g.5553C>T in intron 2 of Bagnolese sheep *DGAT*1 by *BamH*I PCR-RFLP. Lane 1, TT homozygous sample; lane 2, heterozygous sample; lane 3, CC homozygous sample; lane 4, 1kb Opti-DNA Ladder, 0.1–10 kb (Applied Biological Materials, ABM); **(B)** Genotyping of the SNP g.8539C>T in the *DGAT*1 exon 17 by *Msp*I PCR-RFLP. Lane 5, CC homozygous sample; lane 6, heterozygous sample; lane 7, TT homozygous sample; lane 8, 1kb Opti-DNA Ladder, 0.1–10 kb (Applied Biological Materials, ABM).

While for the first mutation, genotyping protocols are not reported in the literature, for the mutation in the 17th exon, various authors apply a genotyping method based on PCR-RFLP using the *Alu*I

endonuclease, whose restriction site (AG/C**T**, bold and underlined the polymorphic site) is altered in the presence of Cytosine. Therefore, individuals with C are undigested.

The choice in this research to create a new protocol that involves restriction in the presence of Cytosine was motivated by the high frequency of this allele in most investigated breeds. The use of the *Msp*1 endonuclease could provide greater assurance for a more accurate and less ambiguous genotyping, especially where a limitation of the *Alu*I restriction method for the rare allele could be hypothesized due to: low enzyme activity, operator error, incorrect maintenance of incubation temperature, pipetting error, etc.

Furthermore, restriction fragments obtained with previous protocols (272 bp and 37 bp *vs* 309 bp) are less easily distinguishable then those proposed in this work (234 bp and 131 bp *vs* 365 bp) (Fig 3S). Therefore, another limitation of the *Alu*I PCR-RFLP method is the suboptimal discrimination of restriction fragments compared to that proposed in this study.

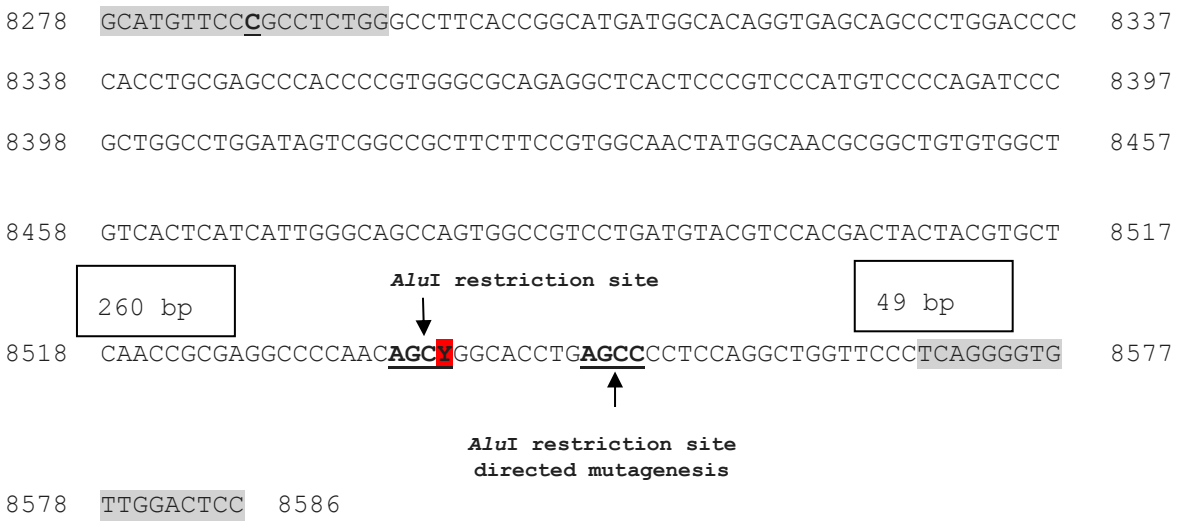

Figure 3S. DNA segment between the 16th and 17th exons of the ovine *DGAT*1 gene, amplified for the optimization of the *Alu*I PCR-RFLP method (Xu *et al*., 2008). The C>T transition occurring at the 147th nucleotide of exon 17 (EU178818.1: g.8539C>T) is highlighted in red. The forward and reverse primers are shaded in gray. The restriction site of the *Alu*I endonuclease (AG/CT) is underlined, along with the restriction site of *AluI* for site directed mutagenesis.

3.2. *Genotyping*

The Bagnolese sheep population under study was genotyped for the g.5553C>T mutation by *BamH*I -PCR-RFLP. The results of the genotype distribution and allele frequencies for this marker are shown in Table 1.

	Genotype Numbers			Allele Frequency		Population Indices		
	CC	CT	TT	C	T	Ho	He	Fis
Obs	90	105	57	0.5655	0.4345	0,4167	0.4924	0.1521
Exp	80,4573	124,0855	47,4573					
χ^2=5,986216								
p=0.014418								
d.f.=1								

Table 1. Genotype numbers (observed and expected), allele frequencies and, population indices observed at the SNP g.5553C>T of *DGAT1 locus* in Bagnolese population (n= 252).

The results of the study carried out on 252 subjects show a higher frequency of the C allele (0.5655). Based on the expected genotypic frequencies, a statistically significant heterozygote deficit is observed (χ2=5.99; $p < 0.05$; d.f.=1). Specifically, the expected frequency of the CT genotype (He) is 0.49, while the observed frequency (Ho) is 0.42. The fixation index (Fis=0.1521), which assesses the level of heterozygosity within a population, confirms the excess of homozygotes. This situation is consistent with the average number of animals per farm (85) and the tendency of farmers to rely mainly on internal breeding, which limits the introduction of breeders from different genetic lines. The g.5553C>T mutation has only been studied to a limited extent. So far, it has been characterized by Scatà *et al*., (2009) in three Italian sheep breeds: Altamurana, Gentile di Puglia and Sarda. According to these authors, all three breeds show a higher frequency of the C allele, especially in the Altamurana and Gentile di Puglia breeds. In particular, a statistically significant association between this SNP and milk fat content was observed for these breeds. Comparing the results of the present study with those reported by Scatà et al., (2009) Bagnolese sheep show a higher frequencies of C allele as the Sarda sheep breed reared in Sardinia (Table 2).

Breed	Allele Frequency		References
	C	T	
Bagnolese	0.565	0.435	This research
Altamurana	0.750	0.250	Scatà *et al*. (2009)
Gentile di Puglia	0.680	0.320	
Sarda	0.590	0.410	

Table 2. Allele frequency of the SNP at position 1415 of intron 2 of the *DGAT*1 gene (g.5553C>T) in the Bagnolese population studied and, in the breeds, studied by Scatà *et al*. (2009).

*Msp*I-PCR-RFLP was used to genotype the g.8539C>T transition at the *DGAT1 locus* of the Bagnolese sheep. The results of the genotype distribution and allele frequencies for this marker are shown in Table 3.

	Genotype Numbers			Allele Frequency		Population Indices		
	CC	CT	TT	C	T	Ho	He	Fis
Obs	196	19	1	0.95	0.05	0.09	0.09	0.049
Exp	195.49	20.03	0.49					
χ^2=0.59								
p=0.44								
d.f.=1								

Table 3. Number of genotypes (observed and expected), allele frequencies and population indices observed for the SNP at position 147 of exon 17 of the *DGAT*1 gene (g.8539C>T) in the studied Bagnolese population (n= 216).

The results of the survey of 216 subjects show a predominant frequency of the C allele (0.95), which may indicate that this is the native allele in the Bagnolese breed. Furthermore, the positive association with improved milk processing characteristics (Dervishi *et al*., 2015), such as lactose content, C4:0 fatty acids, C16:1 c9 and the n-6: n-3 ratio, could have favored the maintenance of the high frequency of C allele in the Bagnolese population by unconscious selection on the part of breeders. Based on the expected and observed genotype frequencies (Ho and He = 0.09), the Hardy-Weinberg equilibrium is respected (χ2=0.59; p=0.44; d.f.=1).

The presence of cytosine at nucleotide 147 of exon 17 characterizes the remaining ruminant species and is therefore considered the ancestral form of the *DGAT*1 gene. However, exceptions to this are ruminants such as the takin (Budorcas taxicolor, GeneBank XM_052651466.1) and the sabre-horned oryx (Oryx dammah, GeneBank XR_005724623.1), which would be characterized by the presence of T.

Although we cannot exclude the possibility that this mutation may have occurred as a result of convergent evolution (allelic lines with similar characteristics evolving independently in separate lines) or genetic introgression (allelic lines being transferred horizontally either from recipient to donor species or in both directions), a possible hypothesis is that we are in the presence of a Trans-Specific Polymorphism (TSP), which refers to the presence of identical or similar alleles in related species. By definition, TSPs originate from the transfer of alleles from an ancestral species to a descendant species by means of 'incomplete lineage sorting' (Klein 1987; Klein *et al*., 1998, 2007).

Common polymorphisms are considered TSPs in the strict sense only if there is convincing evidence that they are lineage-identical mutations rather than recurrent mutations occurring independently in different lines (identical in status). The latter often involve the CpG nucleotides (Klein *et al*., 1998; Leffler *et al*., 2013), whose impermutability is directly attributable to their role as the major methylation site of cytosine, with the consequent risk of spontaneous deamination of 5-methylcytosine to produce thymine.

Two different forms of TSP are generally distinguished: neutral TSP and balanced TSP. Neutral TSP is common in closely related species that have recently diverged and are gradually disappearing

(Klein *et al*., 1998). Thus, neutral TSP tends to be widespread among loci only within a short time after a speciation event (Nagl *et al*., 1998; Samonte *et al*., 2007). In contrast, the balanced TSP is functionally much more important (Klein *et al*., 2007). This type of TSP is a consequence of balanced selection, which is a prerequisite for the maintenance of genetic variability. Balanced TSP is generally long-lived and can be maintained for millions or tens of millions of years (Kamath and Getz, 2011; Aguilar and Garza, 2007). A particular example of the long-term effects of stabilizing selection is reported for artiodactyls (sheep, goats, cattle, etc.), where it appears to be maintained in allelic lineages for over 20 million years (Gutiérrez-Espeleta *et al*., 2001).

The TSP is a key evolutionary mechanism responsible for the adaptive sharing of genetic variation between taxa.

Although most studies of TSP have so far focused on MHC loci (especially MHC I and MHC II) (Klein *et al*, 2007; Asthana *et al*, 2005), the few examples described in other genes involved in immune responses suggest that TSP is a common and general evolutionary phenomenon.

Indeed, in addition to immune system genes, TSP has been documented for self-incompatibility loci that prevent self-fertilisation in Angiosperms (Ioerger *et al*., 1990; Richman *et al*, 1996; Dwyer *et al*., 1991), in 'mating' loci in fungi (Lukens *et al*., 1996; Muirhead *et al*., 2002), in the ABO blood system (Kermarrec *et al*., 1999; Martinko *et al*., 1993; Ségurel *et al*., 2012), in the retroviral restriction factor (*TRIM*5 gene) (Cagliani *et al*., 2010), in the antiviral zinc-finger protein (*ZC3HAV*1 gene) (Cagliani *et al*., 2012) and in the gene encoding Ladinin-1 (Teixeira *et al*., 2015) in primates or in the *CDS* (Complementary Sex Determiner) gene in hymenopterans (Lechner *et al*., 2014; Heimpel and de Boer, 2008) and pufferfish (Kamiya *et al*., 2012).

Recently TSPs have been found and characterized in ruminants at the *OXT* (Cosenza *et al*., 2017) and *PRLR* (Cosenza *et al*., 2018) *loci*.

Therefore, if confirmed in other species, the g.8539C>T polymorphism could be another example of TSPs detected at the DNA level in ruminants, and the first in a gene encoding an enzyme.

The mutation g.8539C>T compared to the transition g.5553C>T, has been the subject of several investigations in many breeds/genetic types reared in different European and non-European nations and much of the research has been aimed at identifying associations with features of interest such as milk and meat traits. Interestingly allelic and genotypic frequency is similar among most of the analyzed breeds (Table 4).

The frequency of the C allele varies from 1 to 0.93 for the breeds reared in Italy and from 1 to 0.75 (Scatà et al., 2009) for those bred in Spain, Romania, Indonesia, Turkey, Egypt, and India.

Similarly, Yang *et al*. (2011) report that the C allele is always predominant in 4 Chinese meat breeds (Tan, Ganjia, Oula and Qiaoke), but with a lower C allele frequency (0.6 to 0.7). However, these

results contrast with those previously reported by Xu *et al*. (2009) for the same (Tan) or other sheep breeds (Small-tailed Han and Inner Mongolia) reared in China, where the frequency ratios between C and T were reversed. Similarly, the allele T is reported to be predominant in the Iranian meat breeds Moghami, Zell and Lori Bakhtiari (Mohammadi *et al*., 2013; Noshahr and Rafat, 2014). The exception is the same Iranian breed Lori, characterized by Nanekarani *et al*. (2016), where an almost equal frequency of the two alleles is found.

The different allele frequencies observed between different sheep breeds could be caused by several factors like the productive attitude (meat, milk or wool), the different geographical area of origin and breeding, the incorrect genotyping due to misinterpretation of results.

Breed	Genotype Frequency			Allele Frequency		References	Country
	CC	CT	TT	C	T		
Bagnolese	0.91	0.088	0.001	0.95	0.05	This study	Italy
Altamurana	-	-	-	0.94	0.06	Scatà *et al*. (2009)	
Gentile di Puglia	-	-	-	0.93	0.07		
Sarda	-	-	-	1	0		
Ansotana	0.74	0.18	0.08	0.83	0.17	Dervishi *et al*. (2015)	Spain
Latxa	0.83	0.17	0.00	0.92	0.08		
Romanov	0.76	0.18	0.06	0.85	0.15		
Rasa aragonesa	0.62	0.33	0.04	0.80	0.20		
Churra	0.79	0.19	0.02	0.88	0.12		
Churra tensina	0.53	0.33	0.14	0.69	0.31		
Churra lebrijana	1.00	0.00	0.00	1	0		
Manchega	0.73	0.21	0.06	0.83	0.17		
Assaf	0.93	0.07	0.00	0.96	0.04		
Compass Agrinac	1.00	0.00	0.00	1	0	Amri *et al*. (2023)	Indonesia
Barbados Cross	1.00	0.00	0.00	1	0		
Jonggol sheep	1.00	0.00	0.00	1	0		
Javanese thin-tailed	0.86	0.13	0.00	0.93	0.06		
Compass Agrinac	1.00	0.00	0.00	1	0	Gunawan *et al*. (2019)	
Barbados Cross	1.00	0.00	0.00	1	0		
Garut Composite	1.00	0.00	0.00	1	0		
Javanese thin-tailed	0.90	0.10	0.00	0.95	0.05		
Javanese fat-tailed	0.94	0.06	0.00	0.99	0.01		
Barki	0.78	0.22	0.00	0.89	0.11	Altwaty *et al*. (2020)	Egypt
Najdi	0.65	0.35	0.00	0.83	0.17		
Harri	0.50	0.50	0.00	0.75	0.25		
Lori	0.433	0.258	0.309	0.562	0.438	Nanekarani *et al*. (2016)	Iran
Lori-Bakhtiari	0.16	0.19	0.65	0.254	0.746	Mohammadi *et al*. (2013)	
Zel	0.07	0.24	0.69	0.19	0.81		
Moghani	0.04	0.262	0.698	0.171	0.829	Noshahr e Rafat (2014)	
Turcana	0.84	0.12	0.04	0.90	0.10	Tăbăran *et al*. (2014)	Romania
Small-tailed Han	0.34	0.09	0.57	0.38	0.62	Xu *et al*. (2009)	China
Tan	0.36	0.11	0.53	0.41	0.59		
InnerMongolia	0.25	0.02	0.73	0.26	0.74		
Tan	0.586	0.379	0.034	0.776	0.224	Yang *et al*. (2011)	
Oula	0.513	0.282	0.205	0.654	0.346		
Ganjia	0.583	0.222	0.194	0.694	0.306		
Qiaoke	0.500	0.235	0.265	0.618	0.382		
Akkaraman	0.91	0.09	0.00	0.96	0.04	Bayram *et al*. (2019)	Turkey
Malpura	0.86	0.02	0.12	0.92	0.08	Meena *et al*. (2016)	India

Table 4. Genotype and allele frequencies of the SNP in position 147 of exon 17 of the *DGAT1* gene (g.8539C>T) in different European and non-European breeds

3.3. *Milk Analyses*

Mean milk yield per lactation (150 days starting after lamb weaning) is 107,00±36,00 Kg with an average daily production of 693,2 ± 233,4 ml/die each ewe. The high milk quantitative production variability may be related to the absence of selection plans for this breed. However, this breed has high milk yield when compared with the average production of the Italian ewes (85 L/ewe) (Pulina et al., 2018).

As regard the quality parameters analyzed, mean fat yield % and Kg are 9,08±1,26 and 9,82±3.24 respectively, while mean protein yield % and Kg are 6,64±0,53 and 7,22±2,11 respectively.

According to these results Bagnolese sheep can be considering a high milk fat breed, showing mean fat yield percentage higher than other Italian autochthonous breeds but also higher than some Italian selected breeds (Scatà et al., 2009; Selvaggi et al., 2017; Cesarani et al., 2023). As regard the protein milk content, Bagnolese sheep align with other Italian native breeds (Scatà et al., 2009).

The milk quality parameters observed in Bagnolese sheep support its breeding that is mainly aimed at Pecorino Bagnolese cheese production, in fact high fat and protein concentrations in the milk are associated with high yields in the resulting dairy products (Bencini and Pulina, 1997).

3.4. SNPs DGAT1 locus effect on milk parameters

All the animals sampled for milk analyses were monomorphic at g.8539C>T of *DGAT1* gene being their genotype CC. As regard the genotype at position 1415 of intron 2 of the *DGAT*1 gene, 29 were CC, 36 were CT and 25 were TT.

Table 5 shows the effects of the genotypes at position 1415 of intron 2 (EU178818.1: g.5553C>T) of *DGAT1 locus* found in the studied animals. No differences in terms of milk yield were observed.

Significant differences among the genotypes were found in fat and protein yield percentage. In detail, animals carrying the CT genotype produced more fat in percentage per milking if compared to the CC and TT ones ($p < 0.01$). Similar results were found for the percentage of protein yield being CT individuals more productive than CC ones ($p < 0.01$).

Parameters	Genotypes		
	CC *n* = 29	CT *n* = 36	TT *n* = 25
Milk yield (Kg/lactation)	112,51±30,01	109,31±42,35	108,86±22,90
Fat yield (%)	8,43±0,73^{B}	9,75±0,98^{A}	8,81±1,24^{B}
Protein yield (%)	5,78±1,83^{B}	6,75±0,36^{A}	6,64±0,34
Fat yield (Kg)	8,71±3,07	10,45±4,0	9,33±2,45
Protein yield (Kg)	6,43±2,14	7,36±2,83	7,07±1,59

A, B = $p < 0.01$

Table 5. Effect of genotypes at position 1415 of intron 2 (EU178818.1: g.5553C>T) at DGAT1 locus on milk yield and composition of 90 Bagnolese sheep.

These results are in line with those reported by Scatà et al. (2009) that observed a positive effect of T allele on fat milk yield percentage in Altamurana and Gentile di Puglia sheep. In addition, Bagnolese sheep also show a positive effect of T allele on protein milk yield percentage as already observed in cattle (Khan et al., 2021).

Interestingly there is an excess of homozygotes indicating that the selection process is undergoing.

Currently, it is unclear the reason why this SNP located on intron 2 of *DGAT1* gene has an effect on fat and protein milk yield percentage but it can be hypothesized that it is associated to other causative SNP that are still unknown.

4. Conclusions

The main advances in sheep dairy industry are observed in those countries where it is supported by scientific research like France and Italy (Pulina *et al.*, 2018). Thus, the availability of sensitive, specific and cheap protocols for genotyping of quantitative loci in sheep will facilitate the optimization productions also in breeds that don't undergo to genetic selection.

The new genotyping technique reported in this paper enables a more specific and cheap characterization of the *DGAT1 locus* in sheep.

Moreover, in this study the autochthonous Bagnolese sheep have been genetically characterized for the first time at *DGAT1 locus* thus increasing the knowledge about the effects of this gene on milk fat production in sheep. Moreover, for the first time in sheep, an effect of this gene on milk protein concentration has been detected.

Since Italian dairy sheep industry is mainly based on local breeds, the data here reported are useful for valorization of niche dairy products like Pecorino Bagnolese, which is the basis of biodiversity safeguard, In addition it will improve the sheep farming incoming thus contrasting the crisis that this livestock sector has been facing for some years.

References

1. Aguilar, Andres, and John Carlos Garza. (2007) Historical balancing selection models on the class II β gene of the major histocompatibility complex in salmonids. Journal of Molecular Evolution 65: 34-43.
2. Altwaty, Nada H., Lamiaa M. Salem, and Karima F. Mahrous. (2020) Single nucleotide polymorphisms in the growth hormone receptor gene and *Alu1* polymorphisms in the diacylglycerol acyltransferase 1 gene in relation to meat production in sheep. Veterinary World 13.5: 884.
3. Amri, F., A. Gunawan and C. Sumantri. (2023) Effect of DGAT1 gene on hot and cold carcass, neck and non-carcass traits in Indonesian sheep. Jurnal Ilmu Produksi dan Teknologi Hasil Peternakan 11.1: 48-53.
4. Asthana, Saurabh, Steffen Schmidt and Shamil Sunyaev. (2005) A limited role for balancing selection. Trends in Genetics 21.1: 30-32.
5. Bayram D, Akyüz B, Arslan K, Özdemir F, Aksel EG, Çinar MU (2019) DGAT1, CAST and IGF-I Gene Polymorphisms in Akkaraman Lambs and Their Effects on Live Weights up to Weaning Age. Kafkas Univ Vet Fak Derg 25 (1): 9-15.
6. Bencini R and Pulina G. (1997) The quality of sheep milk: A review. Animal Production Science 37:485-504.
7. Cagliani R, Fumagalli M, Biasin M, Piacentini L, Riva S, Pozzoli U, Bonaglia MC, Bresolin N, Clerici M, Sironi M. (2010) Long-term balancing selection maintains trans-specific polymorphisms in the human TRIM5 gene. Human Genetics 128: 577-588.
8. Cagliani R, Guerini FR, Fumagalli M, Riva S, Agliardi C, Galimberti D, Pozzoli U, Goris A, Dubois B, Fenoglio C, Forni D, Sanna S, Zara I, Pitzalis M, Zoledziewska M, Cucca F, Marini F, Comi GP, Scarpini E, Bresolin N, Clerici M, Sironi M. (2012) A trans-specific polymorphism in ZC3HAV1 is maintained by long-standing balanced selection and may confer susceptibility to multiple sclerosis." Molecular Biology and Evolution 29.6: 1599-1613.
9. Cases S, Smith SJ, Zheng YW, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ, Erickson SK, Farese RV Jr. (1998) Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. Proc Natl Acad Sci U S A. 1998 Oct 27;95(22):13018-23. doi: 10.1073/pnas.95.22.13018.
10. Cesarani A, Mastrangelo S, Congiu M, Portolano B, Gaspa G, Tolone M, Macciotta NPP. Relationship between inbreeding and milk production traits in two Italian dairy sheep breeds. J Anim Breed Genet. 2023 Jan;140(1):28-38. doi:10.1111/jbg.12741. Epub 2022 Oct 14. PMID: 36239218; PMCID: PMC10092622.
11. Chen HC, Smith SJ, Ladha Z, Jensen DR, Ferreira LD, Pulawa LK, McGuire JG, Pitas RE, Eckel RH, Farese RV Jr. (2002b) Increased insulin and leptin sensitivity in mice lacking acyl CoA:diacylglycerol acyltransferase 1. J Clin Invest. 109(8):1049-55. doi: 10.1172/JCI14672. PMID: 11956242; PMCID: PMC150948.
12. Chen HC, Stone SJ, Zhou P, Buhman KK, Farese RV Jr. (2002a) Dissociation of obesity and impaired glucose disposal in mice overexpressing acyl coenzyme a:diacylglycerol acyltransferase 1 in white adipose tissue. Diabetes. 51(11):3189-95. doi: 10.2337/diabetes.51.11.3189.
13. Coppieters W, Riquet J.S, Arranz J-J., Berzi P., Cambisano N., Grisart B., Karim L., Marcq F., Moreau L., Nezer C., Simon P., Vanmanshoven P., Wagenaar D., Georges M. (1998) A QTL with major effect on milk yield and composition maps to bovine chromosome 14. Mamm. Genome, 9:540-544.
14. Cosenza G, Iannaccone M, Auzino B, Macciotta NPP, Kovitvadhi A, Nicolae I, Pauciullo A. (2018) Remarkable genetic diversity detected in the buffalo prolactin receptor (PRLR) gene and association studies with milk fatty acid composition. Animal Genetics 49.3: 159-168.
15. Cosenza G, Macciotta NPP, Nudda A, Coletta A, Ramunno L, Pauciullo A. (2017) A novel polymorphism in the gene encoding for the oxytocin receptor (OXTR) influences milk fatty acid composition in Italian Mediterranean river buffalo. Journal of Dairy Research 84.2: 170-180.
16. Dai R, Zhou H, Fang Q, Zhou P, Yang Y, Jiang S, Hickford JGH. (2022) Variation in ovine DGAT1 and its association with carcass muscle traits in Southdown sheep. Genes 13.9: 1670.
17. Dervishi, E., Serrano M., Joy M., Sarto P., Somera A., Gonzàlez-Calvo L., Berzal-Herranz B., Molino F., Martinez-Royo A., Calvo J.H. (2015) Structural characterisation of acyl CoA: diacylglycerol acyltransferase 1 (DGAT1) gene and association studies with milk traits in the Assaf sheep breed. Small ruminant research 131: 78-84.
18. Dwyer, Kathy G., et al. (1991) DNA sequences of self-incompatibility genes from Brassica campestris and B. oleracea: polymorphism preceding speciation. Plant Molecular Biology 16: 481-486.
19. Fürbass R, Winter A, Fries R, Kühn C. (2006) Alleles of the bovine DGAT1 variable number of tandem repeat associated with a milk fat QTL at chromosome 14 can stimulate gene expression. Physiol Genomics. 25(1):116-20. doi:10.1152/physiolgenomics.00145.2005.
20. Grisart, Jacques M. and Martial Van der Linden. Conscious and automatic uses of memory in patients with chronic pain. Pain 94.3 (2001): 305-313.
21. Grisart, B.; Farnir, F.; Karim, L.; Cambisano, N.; Kim, J.-J.; Kvasz, A.; Mni, M.; Simon, P.; Frère, J.-M.; Coppieters, W.; Georges, M. Genetic and Functional Confirmation of the Causality of the *DGAT1 K232A* Quantitative Trait Nucleotide in Affecting Milk Yield and Composition. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (8), 2398–2403. https://doi.org/10.1073/pnas.0308518100.

22. Gunawan, A., R. S. Harahap., K. Listyarini., C. Suman-tri. (2019) Identifikasi keragaman gen DGAT1 serta asosiasinya terhadap karakteristik karkas dan sifat pe-ternakan domba. Jurnal Ilmu dan Teknologi Peternakan Tropis. 6(2):259–266.
23. Gutierrez-Espeleta GA, Hedrick PW, Kalinowski ST, Garrigan D, Boyce WM. Is the decline of desert bighorn sheep from infectious disease the result of low MHC variation? Heredity (Edinb). 2001 Apr;86(Pt 4):439-50. doi:10.1046/j.1365-2540.2001.00853.x.
24. Heimpel, George E. and Jetske G. De Boer. (2008) Sex determination in hymenoptera'. Anna. Rev. Entomol. 53: 209-230.
25. Ioerger, Thomas R., Andrew G. Clark and Teh-Hui Kao. (1990) Polymorphism at the self-incompatibility locus in Solanaceae precedes speciation. Proceedings of the National Academy of Sciences 87.24: 9732-9735.
26. Kamath PL, Getz WM. (2011) Adaptive molecular evolution of the Major Histocompatibility Complex genes, DRA and DQA, in the genus Equus. BMC Evol Biol. 11:128. doi: 10.1186/1471-2148-11-128.
27. Kamiya T, Kai W, Tasumi S, Oka A, Matsunaga T, Mizuno N, Fujita M, Suetake H, Suzuki S, Hosoya S, Tohari S, Brenner S, Miyadai T, Venkatesh B, Suzuki Y, Kikuchi K. (2012) A trans-species missense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, Takifugu rubripes (fugu). PLoS genetics 8.7: e1002798.
28. Kermarrec N, Roubinet F, Apoil PA, Blancher A. (1999) Comparison of the O allele sequences of the ABO system of human and non-human primates. Immunogenetics 49: 517-526
29. Khan MZ, Ma Y, Ma J, Xiao J, Liu Y, Liu S, Khan A, Khan IM, Cao Z. (2021) Association of DGAT1 With Cattle, Buffalo, Goat, and Sheep Milk and Meat Production Traits. Front Vet Sci. 16;8:712470. doi:0.3389/fvets.2021.712470.
30. Klein, Jan et al. Trans-species molecular polymorphism. (1998) Annual Review of Ecology and Systematics 29.1: 1-21.
31. Klein, Jan, Akie Sato and Nikolas Nikolaidis. (2007) MHC, TSP and the origin of species: from immunogenetics to evolutionary genetics. Anna. Rev. Genet. 41: 281-304.
32. Klein, Ted M., et al. (1987) High-speed microprojectiles to deliver nucleic acids into living cells. Nature 327.6117: 70-73.
33. Kuehn C, Edel C, Weikard R, Thaller G. (2007) Dominance and parent-of-origin effects of coding and non-coding alleles at the acylCoA-diacylglycerol-acyltransferase (DGAT1) gene on milk production traits in German Holstein cows. BMC Genet. 24;8:62. doi: 10.1186/1471-2156-8-62. PMID: 17892573
34. Kühn C, Thaller G, Winter A, Bininda-Emonds OR, Kaupe B, Erhardt G, Bennewitz J, Schwerin M, Fries R. Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with an important effect on milk fat content in cattle. Genetics 167.4 (2004): 1873-1881.
35. Lechner S, Ferretti L, Schöning C, Kinuthia W, Willemsen D, Hasselmann M. (2014) Nucleotide variability at its limit? Insights into the number and evolutionary dynamics of sex-determining specificities in the honeybee Apis mellifera. Molecular Biology and Evolution 31.2 (2014): 272-287.
36. Leffler, Ellen M., et al. (2013) Multiple examples of ancient balanced selection shared between humans and chimpanzees. Science 339.6127: 1578- 1582.
37. Littell, R.C.; Henry, P.R.; Ammerman, C.B. Statistical analysis of repeated measures data using SAS procedures. J. Anim. Sci. 1998, 76, 1216–1231.
38. Lukens, Lewis, Huang Yicun and Georgiana May. (1996) Correlation of genetic and physical maps in the mating type A locus of Coprinus cinereus. Genetics 144.4: 1471-1477.
39. Mach N, Blum Y, Bannink A, Causeur D, Houee-Bigot M, Lagarrigue S, Smits MA. (2012) Pleiotropic effects of polymorphism of the gene diacylglycerol-O-transferase 1 (DGAT1) in the mammary gland tissue of dairy cows. J Dairy Sci.95(9):4989-5000. doi: 10.3168/jds.2012-5348.
40. Martinko JM, Vincek V, Klein D, Klein J. (1993) Primate ABO glycosyltransferases: evidence for trans-species evolution. Immunogenetics 37: 274-278.
41. Meena A.S.1, Bhatt R.S.1, Sahoo A., Kumar S. (2016) Genetic Polymorphism of The Diacylglycerol Acyltransferase 1 (DGAT1) Gene in Malpura Sheep. Ind J of Small Rum 22:1, 97-99.
42. Mohammadi, Hossein, Mohammad Moradi Shahrebabak and Mostafa Sadeghi. (2013) Association between single nucleotide polymorphism in ovine DGAT1 gene and carcass traits in two Iranian sheep breeds. Animal Biotechnology 24.3:159-167.
43. Muirhead, Christina A., N. Louise Glass and Montgomery Slatkin. (2002) Multilocus self-recognition systems in fungi as a cause of trans-species polymorphism. Genetics 161.2: 633-641.
44. Nagl, Sandra et al. (1998) Persistence of neutral polymorphisms in cichlid fishes from Lake Victoria. Proceedings of the National Academy of Sciences 95.24: 14238-14243.
45. Nanekarani, Shahram, Mojtaba Kolivand and Majid Goodarzi. (2016) Polymorphism of a DGAT1 gene mutation in the Lori sheep breed. J Adv Agri Tech 3.1
46. Nonneman D, Rohrer GA, (2002). Linkage mapping of porcine DGAT1 to a region of chromosome 4 that contains QTL for growth and fatness. Anim Genet 33: 472–473
47. Noshahr FA, Rafat A. (2014). Polymorphism of DGAT1 gene and its relationship with carcass weight and dressing percentage in Moghani sheep breed. Iran J Appl Anim Sci, 4(2): 331-334.

48. Peretti V., Ciotola F., Iannuzzi L. (2013) Characterization, conservation and sustainability of endangered animal breeds in Campania (Southern Italy). Natural Science 5(5A):1-9.
49. Pulina G, Milán MJ, Lavín MP, Theodoridis A, Morin E, Capote J, Thomas DL, Francesconi AHD, Caja G. (2018) Invited review: Current production trends, farm structures, and economics of the dairy sheep and goat sectors. J Dairy Sci. 101(8):6715-6729. doi: 10.3168/jds.2017-14015.
50. Richman, Adam D. and Joshua R. Kohn. (1996) Learning from rejection: the evolutionary biology of single locus incompatibility. Trends in Ecology and Evolution 11.12: 497-502.
51. Samonte, Irene E., et al. (2007) Gene flow among haplochromine fish species from Lake Victoria. Molecular Biology and Evolution 24.9: 2069-2080.
52. Scatà, MC, Napolitano F., Casu S., Carta A., De Matteis G., Signorelli F., Annicchiaro G., Catillo G., Moioli B. (2009) Sheep acyl CoA: molecular characterisation of diacylglycerol acyltransferase 1, polymorphisms and association with milk traits'. Animal Genetics 40.5: 737-742.
53. Schennink A, Heck JM, Bovenhuis H, Visker MH, van Valenberg HJ, van Arendonk JA. Milk fatty acid unsaturation: genetic parameters and effects of stearoyl-CoA desaturase (SCD1) and acyl CoA: diacylglycerol acyltransferase 1 (DGAT1). J Dairy Sci. 2008 May;91(5):2135-43. doi: 10.3168/jds.2007-0825.
54. Ségurel L, Thompson EE, Flutre T, Lovstad J, Venkat A, Margulis SW, Moyse J, Ross S, Gamble K, Sella G, Ober C, Przeworski M. (2012) The ABO blood group is a trans-species polymorphism in primates. Proc Natl Acad Sci U S A 109(45):18493-8. doi: 10.1073/pnas.1210603109.
55. Selvaggi M, D'Alessandro AG, Dario C. Environmental and genetic factors affecting milk yield and quality in three Italian sheep breeds. J Dairy Res. 2017 Feb;84(1):27-31. doi: 10.1017/S0022029916000765. Epub 2016 Dec 23. PMID:28007039.
56. Smith, Lisa C. and Lawrence James Haddad. Overcoming child malnutrition in developing countries: past achievements and future choices . vol. 30. International Food Policy Res Inst, 2000.
57. Tăbăran A., Mihaiu M., Dan SD., Reget O., Pvariu B., Cordis I., Cordea D., Muresan C. (2014) Identification of polymorphism in goat and sheep DGAT1 gene associated with milk production traits. UASVM Bulletin Veterinary Medicine 71.2: 281-286.
58. Teixeira JC, de Filippo C, Weihmann A, Meneu JR, Racimo F, Dannemann M, Nickel B, Fischer A, Halbwax M, Andre C, Atencia R, Meyer M, Parra G, Pääbo S, Andrés AM. (2015) Long-Term Balancing Selection in LAD1 Maintains a Missense Trans-Species Polymorphism in Humans, Chimpanzees, and Bonobos. Mol Biol Evol. 32(5):1186-96. doi: 10.1093/molbev/msv007.
59. Thaller G, Krämer W, Winter A, Kaupe B, Erhardt G, Fries R. (2003) Effects of DGAT1 variants on milk production traits in German cattle breeds. J Anim Sci.81(8):1911-8. doi: 10.2527/2003.8181911x.
60. Winter, A., Krämer, W., Werner, FA, Kollers, S., Kata, S., Durstewitz, G., … & Fries, R. (2002). Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation in a quantitative trait locus for milk fat content. Proceedings of the National Academy of Sciences, 99 (14), 9300-9305.
61. Xu Q.L., Chen YL., Ma RX., Peng X. (2009) Polymorphism of DGAT1 associated with fat-mediated intramuscular soreness in sheep'. Journal of Food Science and Agriculture 89.2: 232-237.
62. Yang Jt, Zang RX, Liu Wj, Xu HW, Bai JL, Lu JX, WU JP (2011) Polymorphism of a Mutation of DGAT1 Gene in Four Chinese Indigenous Sheep Breeds. Asian J of Anim and Vet Adv 6:5, 460-468.
63. Yeh, F.C.; Yang, R.; Boyle, T.J.; Ye, Z.; Xiyan, J.M. PopGene32, Microsoft Windows-based freeware for population genetic analysis; Version 1.32; University of Alberta: Edmonton, Canada, 2000.
64. Yuan J., Zhou J., Deng X., Hu Z., Li N. (2007) Molecular cloning and detection of single nucleotide polymorphism of the buffalo DGAT1 gene. Biochemical Genetics 45: 611-621.
65. Zhang, Q.; Boichard, D.; Hoeschele, I.; Ernst, C.; Eggen, A.; Murkve, B.; Pfister-Genskow, M.; Witte, L. A.; Grignola, F. E.; Uimari, P.; Thaller, G.; Bishop, M. D. Mapping Quantitative Trait Loci for Milk Production and Health of Dairy Cattle in a Large Outbred Pedigree. Genetics 1998, 149 (4), 1959–1973. https://doi.org/10.1093/genetics/149.4.1959.

CHAPTER 3

Characterization of *MBL2* gene and association study with milk Somatic Cell Count (SCC) in Italian Mediterranean River Buffalo (Bubalus bubalis)

1. Introduction

The Italian Mediterranean River Buffalo (IMRB) is a precious heritage of Italian agriculture, both for its historical importance and for its current role in the dairy industry. This buffalo breed, officially recognized by Ministry of Agriculture and Forestry (MIPAF) in 2000, has been selected over the years with clearly defined objectives. Bred mainly to produce milk for the well-known Mozzarella di Bufala Campana PDO, the IMRB has undergone continuous genetic evolution to improve both the quantity and quality of the milk production. The main aim of genetic selection is to increase the production of mozzarella while guaranteeing the quality of the milk. This means not only increasing the quantity of milk produced, but also improving its organoleptic and nutritional characteristics. In addition, constant efforts are being made to improve resistance to technical diseases and the overall resistance of the breed. These efforts not only contribute to the sustainability of buffalo farming, but also to the preservation of the tradition and authenticity of Italian dairy products. Genetic selection is a fundamental tool to improve resistance and resilience traits in buffalo breeds such as the Italian Mediterranean River Buffalo. Research efforts are focused on identifying and studying genetic variants that confer economic benefits in terms of animal health and resilience. In this context, institutions play a crucial role in promoting research and identifying promising genes that influence resistance to infectious diseases and overall resiliency.

MBL gene (Mannose-Binding Lectin) has been shown to be particularly promising in this context. This gene, encoding for mannose-binding lectin (MBL), is an acute-phase protein that binds to mannose and N-acetyl-D-glucosamine sugars present on a wide range of pathogens, including gram-positive and gram-negative bacteria, yeasts, parasites, and viruses (Turner 1998; Takahashi et al. 2005). This binding on microbial surface may cause the direct absorption of the microbial agent from the phagocytes or the activation of a pro-serin proteases complex formed by MASP 1 e MASP 2 leading to the splitting of C4 and C2 of the complement system. (Thiel et al. 1997, Nepomuceno et al. 1997, Turner 1998). The antimicrobial functions of MBL depend on the correct oligomeric assembly and clustered orientation of carbohydrate recognition domains. These domains recognize and bind to molecular patterns associated with pathogens repeated on microbial surfaces (Van de Wetering et al., 2004).

Most mammalian species have *MBL1* and *MBL2* genes, encoding for MBL-A and MBL-C proteins, respectively. Studies have highlighted some mutations in both genes contributing to variation in susceptibility to various infections (Capparelli et al., 2008, Lilli et al., 2005, Shi et al., 2004).
Although the *MBL2* gene has always been considered to consist of 4 exons and 3 introns as described by Gjerstorff et al. (2012), the sequence published in Genebank for the Murrah buffalo consists of 5 exons and 4 introns, although exon number 1 is non-coding (https://www.ncbi.nlm.nih.gov/gene/?term=mbl2%20gene%20buffalo).
Genetic variability studies of the *MBL2* gene in ruminants are rather limited. In Italian Mediterranean River Buffalo, Capparelli et al. (2008) studied gene variants associated with resistance to *Brucella abortus* and Shregojry et al. (2023) studied the gene in relation to mastitis, identifying 4 SNPs potentially associated with mastitis risk.
Therefore, based on these considerations, in this study *MBL2* gene was characterized in all its exonic regions as well as the promoter region in the Italian Mediterranean River Buffalo, for the first time. In addition, it has been checked if there were SNPs associated with milk quality parameters, given the previous discoveries in other ruminants (Wang et al., 2012, Zhao et al., 2012, Shregojry et al., 2023)
The integration of *MBL2* gene variants into IMRB breeding programs could allow more efficient selection for improved mammary health and disease resistance, ultimately leading to improved productivity, sustainability and profitability of dairy farms.

2. Materials and Methods

Animals

Blood samples were collected from 121 lactating Italian Mediterranean River Buffaloes from 7 herds, located in Lazio and Campania regions. The blood samples were collected with anticoagulant EDTA (K3) and stored at 4°C until extraction. Information on the Somatic Cell Count (SCC) of the milk has been provided by the ANASB (Associazione Nazionale Allevatori Specie Bufalina - National Buffalo Species Breeders Association) as a result of the periodic records used to supervise the production.

Genomic DNA extraction, PCR amplification and Sequencing

Genomic DNA was extracted from peripheral blood samples using the Wizard Genomic DNA Purification kit (Promega Corporation) according to the protocol of the manufacturer. To assess the quality and quantity of DNA, the samples were measured using a bio-photometer. Seven primer pairs were designed by using Primer3 software based on the Murrah Buffalo's published sequence (GenBank accession number NC_059179, region. 6688589..6695693) to amplify the 5 exons and part

of the introns near them, as well as the promoter and 3'UTR regions. Sequence of primers, target region and amplicon sizes are provided in Table 1.

The PCR mix contained 1.25U of Taq G2 DNA polymerase (Promega), 200μM for each dNTP (Promega), 100-150ng of DNA template, 0.1 μM of each primer (Invitrogen), with the addition of MilliQ water up to 20μl. The amplification was performed using a BIO-RAD MyCycler thermocycler with the following conditions: pre denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 40 s, annealing at 60°C for 40 s for each set of primer and extension at 72°C for 40 s, and a final extension at 72°C for 10 min. 5 μL of PCR product was electrophoresed on 1.5% agarose gel to check the quality and specificity of amplified DNA product. The PCR products obtained after optimization of PCR conditions were 611bp, 700bp, 504bp, 503bp, 278bp, 582bp and 607bp for primers pairs 1, 2, 3, 4, 5, 6 and 7 respectively. The amplicons were sequenced using Sanger method in Applied Biosystem 3130 Genetic Analyzer. The resulting sequences were analyzed with the DNASTAR SeqMan programme for characterization and identification of SNPs in the target regions of the gene. The sequence of the amplicons was compared with the reference sequence of the Murrah buffalo.

Statistical Analysis

Genotypic frequencies, allelic frequencies, polymorphism information content (PIC), heterozygosity (He) and Hardy-Weinberg equilibrium (HWE) were analyzed using POPGENE32 (ver.1.31) (Yeh et al., 1999). The data underwent a preliminary processing phase, which included the removal of missing data and the handling of outliers, in order to ensure the quality and reliability of the analysis. A χ^2 test was then performed to assess the existence of a significant relationship between the variables considered, i.e. phenotypic production data and genotype. To calculate the correlations between the SNP (categorical) variables and the indicated quantitative phenotypic variables, analysis of variance (ANOVA) was used for each combination of SNP and phenotypic variable. ANOVA allowed us to detect significant differences in the mean phenotypic variables between the different groups defined by the SNP variables.

Gene region	Primer	Sequence (5'- 3')	Fragment length (bp)	Annealing T (°C)
Promoter region	760_bb_MBLprom_F	CACTACAGGTTGGGGCTGAC	611bp	60°
	761_bb_MBLprom_R	TGCACAATTCCCACTTCAAA		
Exon1-5'UTR	762_bb_MBL_e1_F	GTCAAAAGCAGCAGGGTGTT	700bp	60°
	763_bb_MBL_e1_R	TCCCTCTTGGCTCTCTTTGA		
Exon2	744_bb_MBL_e2_F	CCAAATTCCCAGTTTGGAGA	504bp	60°
	745_bb_MBL_e2_R	TGTGTGAGAGTCTGGGATGC		
Exon3	746_bb_MBL_e3_F	GCATCCCAGACTCTCACACA	503bp	60°
	747_bb_MBL_e3_R	GAGAAGCTGGAGGCTGCTTA		
Exon4	748_bb_MBL_e4_F	CAAAAGGTACCCTACCCCCTTA	278bp	60°
	749_bb_MBL_e4_R	CCAAAAGTTGTTGTGAGGATTTC		
Exon5-1	750_bb_MBL_e5-1_F	CATACCATATTTTATGTGGCATCTG	582bp	60°
	751_bb_MBL_e5-1_R	TCAGTACGGGGTGTGTGAAA		
3'UTR-3'Flanking region	750_bb_MBL_e5-2.2_F	ACTTTTGTTGGACGGCAAAT	607bp	60°
	751_bb_MBL_e5-2.2_R	TTCCAAAGAAATCCCAGGAC		

Table 1: Sequences of the seven primer pairs designed to characterize *MBL2* gene in IMRB.

3. Results and Discussions

MBL2 gene characterization

In this work, the *MBL2* gene was characterized for the first time in the Italian Mediterranean River Buffalo. The characterization was carried out by Sanger sequencing, which revealed a total of 33 SNPs compared to the reference sequence of the Murrah buffalo.

The SNPs were distributed as follows: 12 in the five exones, 8 in the promoter region, 8 in the 3'UTR and 3'flanking regions and 5 in the intronic regions (Table 2).

Nine SNPs were found in the coding region, seven of which were non-synonymous and were located in exons 2, 3 and 5. These mutations occur at nucleotide positions 6690911, 6690949, 6690950, 6691373, 6691397, 6694057 and 6694167 (accession number NC_059179). The resulting amino acid substitutions are showed in Table 2.

Furthermore, if the SNPs at position 66909949 and 6690950 occur together (CGC-->UAC), they lead to an additional amino acid substitution Arginine (Arg)→Tyrosine (Tyr) at position 53 of the MBL2 protein. The other SNPs are located in the non-coding regions of the gene, i.e. intronic regions and the 3' and 5' UTR regions. Further analysis of the untranslated sequences is needed to see whether these SNPs could somehow alter the splicing sites or modify the ORFs, which could then lead to alternative splicing and thus the production of structurally different proteins.

Assuming that the sequence identified as exon 1 of the buffalo *MBL2* gene belongs to the leader peptide, the SNPs found in this region are of particular interest, as it has been shown that variations in the leader peptide could influence the correct synthesis of the protein (Kozak 2005). In multicellular eukaryotes, translation is known to be influenced by various aspects of mRNA structure, including the m7G cap, the primary sequence or context surrounding AUG codons, the position of AUG codons (i.e., whether it is the first AUG in the message), leader length, and secondary structure, both upstream and downstream of AUG codons (Kozak 1991). In the present study, three SNPs of exon 1 (6689614T→C, 6689632G→A, 6689665C→T) were identified. Therefore, in light of the above considerations, it is reasonable to assume that these three mutations may play a role in the differential translational efficiency of the *MBL2* gene in buffalo.

Mutations in the promoter region are also of particular interest, as it has been shown in humans that mutations in this region at positions -550 and -221 affect serum protein levels (Madsen et al.,1998, Steffensen et al., 2000, Zhang et al.,2005). However, it should be noted that the above authors described a gene composed of 4 exons and 3 introns, whereas current references indicate the presence of 5 exons and 4 introns. The first of these exons is non-coding, as a result, these considerations lead us to state that variants previously considered to be located in the promoter region are now located in the first intron. This does not make the SNPs found in these regions any less important, as they may

also be involved in the regulation of expression. As for the SNPs along the coding regions, they are mostly non-synonymous, in fact only two out of 9 SNPs result in the conservation of the encoded amino acid (6691392G→A; 6694193C→T) and are in exons 3 and 5, respectively.

The other seven SNPs are responsible for variations in the translation of amino acids, but the most interesting of these are those that determine translation into an amino acid with different chemical properties (Gly40→Asp; Arg53→Cys; Asp64→Gly; Asp72→Gly; Thr167→Met), because differences in these properties of the amino acids can alter the protein architecture and activity.

Moreover, comparing our findings with those from Shregojry et al. (2023), it's clear that 3 out of 4 SNP found by these authors are the same as 3 SNPs found in this work (g.6690911G>A, g.6694057C>T, g.6694193C>T).

N°	***Bp (NC_059179)***	***Alleles***	***Protein Position***	***Codon Changinng***	***Amino Acids***	***Genomic region***
SNP 1	6689015	G/C	-	-	-	5'Flanking region
SNP 2	6689061	T/C	-	-	-	5'Flanking region
SNP 3	6689062	G/A	-	-	-	5'Flanking region
SNP 4	6689080	T/A	-	-	-	5'Flanking region
SNP 5	6689100	G/A	-	-	-	5'Flanking region
SNP 6	6689277	C/T	-	-	-	5'Flanking region
SNP 7	6689336	C/T	-	-	-	5'Flanking region
SNP 8	6689351	A/G	-	-	-	5'Flanking region
SNP 9	6689614	T/C	-	-	-	Exon 1 -5'UTR
SNP 10	6689632	G/A	-	-	-	Exon 1 -5'UTR
SNP 11	6689665	C/T	-	-	-	Exon 1 -5'UTR
SNP 12	6690724	G/T	-	-	-	Intron 1
SNP 13	6690757	C/A	-	-	-	Intron 1
SNP 14	6690911	G/A	p.40	GGU→GAU	Gly→Asp	Exon 2
SNP 15	6690949	C/T	p.53	CGC→UGC	Arg→Cys	Exon 2
SNP 16	6690950	G/A	p.53	CGC→CAC	Arg→His	Exon 2
SNP 17	6691003	T/C	-	-	-	Intron 2
SNP 18	6691190	G/A	-	-	-	Intron 2
SNP 19	6691373	A/G	p.64	GAU→GGU	Asp→Gly	Exon 3
SNP 20	6691392	G/A	p.70	UCG→UCA	Ser→Ser	Exon 3
SNP 21	6691397	A/G	p.72	GAC→GGC	Asp→Gly	Exon 3
SNP 22	6691513	G/A	-	-	-	Intron 3
SNP 23	6694057	C/T	p.167	ACG→AUG	Thr→Met	Exon 5
SNP 24	6694167	G/A	p.204	GUG→AUG	Val→Met	Exon 5
SNP 25	6694193	C/T	p.220	GGC→GGU	Gly→Gly	Exon 5

SNP 26	6694422	T/C	-	-	-	3'UTR
SNP 27	6694524	C/T	-	-	-	3'UTR
SNP 28	6694542	T/C	-	-	-	3'UTR
SNP 29	6694548	A/G	-	-	-	3'UTR
SNP 30	6694573	C/T	-	-	-	3'UTR
SNP 31	6694769	A/G	-	-	-	3' Flanking
SNP 32	6694783	T/C	-	-	-	3' Flanking
SNP 33	6694795	G/A	-	-	-	3' Flanking

Table 2: SNPs of the *MBL2* gene in Italian Mediterranean river buffalo resulting from Sanger sequencing.

Additionally, Shregojry et al. (2023), demonstrated that the alleles A, C and C of the g.6690911G>A, g.6694057C>T, g.6694193C>T SNPs were more frequently represented in animals affected by mastitis as compared to unaffected ones and they had respectively 3.7, 7.82 and 3.7 folds increased risk for developing the clinical mastitis in Murrah buffaloes. As the Italian Mediterranean River Buffalo also showed the same SNPs, it is likely that the variants at these loci are also responsible for resistance to technopathies such as mastitis in this breed.

Scientific evidence confirms the connection between genetics and health status, supported by studies in different species. Wang et al. (2012) and Zhao et al. (2012) identified two SNPs in exon 1 (g.201 G > A, g.234 C > A according to Wang et al., 2012; g.1164 G>A and g.1197 C>A according to Zhao et al., 2012) where individuals homozygous for the wild type allele had lower SCS than heterozygotes and homozygotes for the mutant allele. In addition, Lillie et al. (2007) showed that a mutation in the promoter region of the *MBL2* gene in pigs correlated with reduced hepatic expression of MBL-C. In humans, three SNPs in exon 1 have been identified that result in a functional deficit of MBL, affecting the assembly and stability of the MBL-C protein (Sumiya et al., 1991; Lipscombe et al., 1992; Madsen et al., 1994). Three other SNPs in the *MBL2* gene promoter affect MBL levels (Beltrame et al., 2015). The combination of these SNPs increases susceptibility to various bacterial, viral, and parasitic diseases (Eisen and Minchinton, 2003; Holmskov et al., 2003).

Table 3 shows population indices of all the 33 SNPs found in the analyzed animals. All SNPs found in the tested population were in Hardy-Weinberg equilibrium. Fis parameter indicates a significant genetic differentiation.

N°	*Bp (NC_059179)*	*Alleles*	*He*	*Ho*	*Fis*	χ^2	*df*
SNP 1	6689015	G/C	0,152	0,115	0.2369	0.172	1
SNP 2	6689061	T/C	0,152	0,115	0.2369	0.172	1
SNP 3	6689062	G/A	0,152	0,115	0.2369	0.172	1
SNP 4	6689080	T/A	0,152	0,115	0.2369	0.172	1
SNP 5	6689100	G/A	0,152	0,115	0.2369	0.172	1
SNP 6	6689277	C/T	0,159	0,107	0.3221	0.172	1
SNP 7	6689336	C/T	0,439	0,380	0.1298	0.928	1
SNP 8	6689351	A/G	0,152	0,115	0.2369	0.172	1
SNP 9	6689614	T/C	0,159	0,107	0.3221	0.172	1
SNP 10	6689632	G/A	0,361	0,305	0.1509	0.623	1
SNP 11	6689665	C/T	0,439	0,380	0.1298	0.928	1
SNP 12	6690724	G/T	0,145	0,107	0.2575	0.332	1
SNP 13	6690757	C/A	0,152	0,115	0.2369	0.172	1
SNP 14	6690911	G/A	0,405	0,380	0.0592	0.688	1
SNP 15	6690949	C/T	0,152	0,115	0.2369	0.172	1
SNP 16	6690950	G/A	0,079	0,082	-0.0431	0.159	1
SNP 17	6691003	T/C	0,152	0,115	0.2369	0.172	1
SNP 18	6691190	G/A	0,462	0,454	0.0130	1.2	1
SNP 19	6691373	A/G	0,475	0,421	0.1093	0.255	1
SNP 20	6691392	G/A	0,462	0,454	0.0130	1.2	1
SNP 21	6691397	A/G	0,152	0,115	0.2369	0.172	1
SNP 22	6691513	G/A	0,462	0,454	0.0130	1.2	1
SNP 23	6694057	C/T	0,484	0,479	0.0054	0.013	1
SNP 24	6694167	G/A	0,008	0,008	-0.0041	0.38	1
SNP 25	6694193	C/T	0,366	0,380	-0.0431	0.012	1
SNP 26	6694422	T/C	0,192	0,215	-0.1204	0.787	1
SNP 27	6694524	C/T	0,102	0,090	0.1058	0.828	1
SNP 28	6694542	T/C	0,457	0,454	0.0026	0.307	1
SNP 29	6694548	A/G	0,192	0,215	-0.1204	0.787	1
SNP 30	6694573	C/T	0,271	0,305	-0.1310	0.031	1
SNP 31	6694769	A/G	0,192	0,215	-0.1204	0.787	1
SNP 32	6694783	T/C	0,192	0,215	-0.1204	0.787	1
SNP 33	6694795	G/A	0,192	0,215	-0.1204	0.787	1

Table 3. Population indices of the 33 SNPs found in the IMRB.

A preliminary χ^2 test has been carried out to verify if there is a significant relationship between SCC and genetic data. This analysis showed that the combinations of SNP 1 - SNP 2, SNP 2 - SNP 17, SNP 15 - SNP 17, SNP 11 - SNP 24, SNP 20 - SNP 24, SNP 18 - SNP 24 have a statistically significant association, with a p-value below the standard significance level of 0.05. These results suggest that many of the SNP variables in the dataset are not independent. The presence of significant relationships between SNP variables may reflect genetic or inheritance patterns like it happen in the case of haplotypes that will be the subject of further in-depth study, especially if these variables are related to specific traits or characteristics of interest.

This evidence clearly indicates that the gene plays a significant role in the health status of both humans and other animal species, including ruminants, by also affecting udder health, suggesting that although an analysis for the presence/absence of mastitis was not performed in this study, it is likely that there is a relationship between gene variants and resistance to infectious diseases.

4. Conclusions

The primary aim of this study was to investigate the *MBL2* gene in the Italian Mediterranean River Buffalo population and to assess its association with Somatic Cell Count (SCC). The gene was found to be highly polymorphic, with a total of 33 SNPs detected. Hardy-Weinberg equilibrium analysis indicated that there is no selective pressure for any allele, probably because the introduction of selection criteria based on udder health and general animal health is relatively recent.

These preliminary results, together with the evidence already available in the scientific literature on buffalo (Capparelli et al., 2008, Shregojry et al., 2023) and on other species (Beltrame et al., 2015, Lillie et al., 2005), suggest that the gene may also have an impact on infectious diseases and general health in the Italian Mediterranean River Buffalo population. For this reason, future studies will focus on the identification of the haplotypes present in the population, their distribution, and their association with production parameters.

The identification of allelic variants associated with beneficial production traits may provide valuable information for the selection of the most robust and infection-resistant animals. In conclusion, the use of natural genetic resistance as a tool for the control of infectious diseases in livestock is a promising approach that can be integrated with existing strategies to improve animal health and welfare and ensure the sustainability of livestock production.

References

1. Beltrame, M. H.; Catarino, S. J.; Goeldner, I.; Boldt, A. B. W.; De Messias-Reason, I. J. The Lectin Pathway of Complement and Rheumatic Heart Disease. *Front. Pediatr.* **2015**, *2*. https://doi.org/10.3389/fped.2014.00148.
2. Capparelli, R.; Parlato, M.; Amoroso, M. G.; Roperto, S.; Marabelli, R.; Roperto, F.; Iannelli, D. Mannose-Binding Lectin Haplotypes Influence Brucella Abortus Infection in the Water Buffalo (Bubalus Bubalis). *Immunogenetics* **2008**, *60* (3–4), 157–165. https://doi.org/10.1007/s00251-008-0284-4.
3. Eisen, D. P.; Minchinton, R. M. Impact of Mannose-Binding Lectin on Susceptibility to Infectious Diseases. *Clinical Infectious Diseases* **2003**, *37* (11), 1496–1505. https://doi.org/10.1086/379324.
4. Gjerstorff, M.; Hansen, S.; Jensen, B.; Dueholm, B.; Horn, P.; Bendixen, C.; Holmskov, U. The Genes Encoding Bovine SP-A, SP-D, MBL-A, Conglutinin, CL-43 and CL-46 Form a Distinct Collectin Locus on Bos Taurus Chromosome 28 (BTA28) at Position q.1.8–1.9. Animal Genetics 2004, 35 (4), 333–337. https://doi.org/10.1111/j.1365-2052.2004.01167.x.
5. Holmskov, U.; Thiel, S.; Jensenius, J. C. Collectins and Ficolins: Humoral Lectins of the Innate Immune Defense. *Annu. Rev. Immunol.* **2003**, *21* (1), 547–578. https://doi.org/10.1146/annurev.immunol.21.120601.140954.
6. Kozak, M. An Analysis of Vertebrate mRNA Sequences: Intimations of Translational Control. *The Journal of cell biology* **1991**, *115* (4), 887–903. https://doi.org/10.1083/jcb.115.4.887.
7. Kozak, M. Regulation of Translation via mRNA Structure in Prokaryotes and Eukaryotes. *Gene* **2005**, *361*, 13–37. https://doi.org/10.1016/j.gene.2005.06.037.
8. Lillie, B. N.; Brooks, A. S.; Keirstead, N. D.; Hayes, M. A. Comparative Genetics and Innate Immune Functions of Collagenous Lectins in Animals. *Veterinary Immunology and Immunopathology* **2005**, *108* (1–2), 97–110. https://doi.org/10.1016/j.vetimm.2005.07.001.
9. Lillie, B. N.; Keirstead, N. D.; James Squires, E.; Anthony Hayes, M. Gene Polymorphisms Associated with Reduced Hepatic Expression of Porcine Mannan-Binding Lectin C. *Developmental & Comparative Immunology* **2007**, *31* (8), 830–846. https://doi.org/10.1016/j.dci.2006.11.002.
10. Lipscombe, R. J.; Sumiya, M.; Hill, A. V. S.; Lau, Y. L.; Levinsky, R. J.; Summerfield, J. A.; Turner, M. W. High Frequencies in African and Non-African Populations of Independent Mutations in the Mannose Binding Protein Gene. *Human Molecular Genetics* **1992**, *1* (9), 709–715. https://doi.org/10.1093/hmg/1.9.709.
11. Madsen, H. O.; Garred, P.; Kurtzhals, J. A. L.; Lamm, L. U.; Ryder, L. P.; Thiel, S.; Svejgaard, A. A New Frequent Allele Is the Missing Link in the Structural Polymorphism of the Human Mannan-Binding Protein. *Immunogenetics* **1994**, *40* (1), 37–44. https://doi.org/10.1007/BF00163962.
12. Madsen, H. O.; Satz, M. L.; Hogh, B.; Svejgaard, A.; Garred, P. Different Molecular Events Result in Low Protein Levels of Mannan-Binding Lectin in Populations from Southeast Africa and South America. *The Journal of Immunology* **1998**, *161* (6), 3169–3175. https://doi.org/10.4049/jimmunol.161.6.3169.
13. Nepomuceno, R. R.; Henschen-Edman, A. H.; Burgess, W. H.; Tenner, A. J. cDNA Cloning and Primary Structure Analysis of C1qRP, the Human C1q/MBL/SPA Receptor That Mediates Enhanced Phagocytosis In Vitro. *Immunity* **1997**, *6* (2), 119–129. https://doi.org/10.1016/S1074-7613(00)80419-7.
14. Shergojry, S. A.; Verma, A.; Ghani, M.; Gupta, I. D.; Mir, N. A. Identification of Genetic Polymorphism of the MBL2 Gene and Its Association with Clinical Mastitis in Murrah Buffaloes. *J Genet* **2023**, *102* , 21. https://doi.org/10.1007/s12041-023-01419-9.
15. Shi, L.; Takahashi, K.; Dundee, J.; Shahroor-Karni, S.; Thiel, S.; Jensenius, J. C.; Gad, F.; Hamblin, M. R.; Sastry, K. N.; Ezekowitz, R. A. B. Mannose-Binding Lectin-Deficient Mice Are Susceptible to Infection with *Staphylococcus Aureus*. *The Journal of Experimental Medicine* **2004**, *199* (10), 1379–1390. https://doi.org/10.1084/jem.20032207.
16. Steffensen, R.; Thiel, S.; Varming, K.; Jersild, C.; Jensenius, J. C. Detection of Structural Gene Mutations and Promoter Polymorphisms in the Mannan-Binding Lectin (MBL) Gene by Polymerase Chain Reaction with Sequence-Specific Primers. *Journal of Immunological Methods* **2000**, *241* (1–2), 33–42. https://doi.org/10.1016/S0022-1759(00)00198-8.
17. Sumiya, M.; Tabona, P.; Arai, T.; Summerfield, J. A.; Super, M.; Levinsky, R. J.; Turner, M. W. Molecular Basis of Opsonic Defect in Immunodeficient Children. *The Lancet* **1991**, *337* (8757), 1569–1570. https://doi.org/10.1016/0140-6736(91)93263-9.
18. Takahashi, K.; Shi, L.; Gowda, L. D.; Ezekowitz, R. A. B. Relative Roles of Complement Factor 3 and Mannose-Binding Lectin in Host Defense against Infection. *Infect Immun* **2005**, *73* (12), 8188–8193. https://doi.org/10.1128/IAI.73.12.8188-8193.2005.
19. Thiel, S.; Vorup-Jensen, T.; Stover, C. M.; Schwaeble, W.; Laursen, S. B.; Poulsen, K.; Willis, A. C.; Eggleton, P.; Hansen, S.; Holmskov, U.; Reid, K. B. M.; Jensenius, J. C. A Second Serine Protease Associated with Mannan-Binding Lectin That Activates Complement. *Nature* **1997**, *386* (6624), 506–510. https://doi.org/10.1038/386506a0.
20. Turner, M. W. Mannose-Binding Lectin (MBL) in Health and Disease. *Immunobiology* **1998**, *199* (2), 327–339. https://doi.org/10.1016/S0171-2985(98)80037-5.
21. Van De Wetering, J. K.; Van Golde, L. M. G.; Batenburg, J. J. Collectins: Players of the Innate Immune System. *European Journal of Biochemistry* **2004**, *271* (7), 1229–1249. https://doi.org/10.1111/j.1432-1033.2004.04040.x.

22. Wang, X.; Ju, Z.; Huang, J.; Hou, M.; Zhou, L.; Qi, C.; Zhang, Y.; Gao, Q.; Pan, Q.; Li, G.; Zhong, J.; Wang, C. The Relationship between the Variants of the Bovine MBL2 Gene and Milk Production Traits, Mastitis, Serum MBL-C Levels and Complement Activity. *Veterinary Immunology and Immunopathology* **2012**, *148* (3–4), 311–319. https://doi.org/10.1016/j.vetimm.2012.06.017.
23. Yeh, F.C., Yang, R.C. and Boyle, T. (1999) POPGENE Version 1.32: Microsoft Window-Based Freeware for Population Genetics Analysis. University of Alberta, Edmonton.
24. Zhang, H.; Zhou, G.; Zhi, L.; Yang, H.; Zhai, Y.; Dong, X.; Zhang, X.; Gao, X.; Zhu, Y.; He, F. Association between Mannose-Binding Lectin Gene Polymorphisms and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Infection. *The Journal of Infectious Diseases* **2005**, *192* (8), 1355–1361. https://doi.org/10.1086/491479.
25. Zhao, Z. L.; Wang, C. F.; Li, Q. L.; Ju, Z. H.; Huang, J. M.; Li, J. B.; Zhong, J. F.; Zhang, J. B. Novel SNPs of the Mannan-Binding Lectin 2 Gene and Their Association with Production Traits in Chinese Holsteins. Genet. Mol. Res. **2012**, *11* (4), 3744–3754. https://doi.org/10.4238/2012.October.15.6.

CHAPTER 4

Identification of *MBL2* gene polymorphisms and their association with mastitis susceptibility in Comisana Sheep breed

1. Introduction

The term mastitis defines the inflammation of the mammary gland parenchyma and is the technopathy causing the most economic losses in the dairy industry (Petrovski et al 2006, Halasa et al 2007, Bhosale et al 2014). The agents that can cause mastitis are not only bacterial, but also chemical, mechanical, or thermal. This disease not only affects the health of the animal, but also the characteristics of the milk it produces, in fact mastitis decreases milk yield and modifies its composition as a result of the damage caused in the secretory tissue of the mammary gland (Martí-De Olives et al 2020). Changes in the milk include a reduction in potassium and lactoferrin, a reduction in casein and, because of the lower casein content, a reduction in calcium. In addition, milk from animals with mastitis has a higher Somatic Cell Count (SCC) (Bhosale et al 2014, Martí-De Olives et al 2020). So, a high SCC in milk is associated with a reduced casein/protein ratio and, in general, lower milk production (Petrovski et al 2006, Martí-De Olives et al 2020). In sheep the literature reports individual milk yield losses of 2.6–43.1% (Libera et al., 2021) depending on infection severity and causative microorganism (Martí-De Olives et al., 2020).

Unfortunately, there is no information on estimated economic losses for sheep, but in cows Petrovsky et al. (2006), estimated that the economic losses associated with this condition ranged from $161.79 to $344.16 per lactating cow per year. These losses are mainly due to reduced milk production, treatment costs and the need for culling (Halasa et al., 2007).

Given the critical role of udder health in dairy animals and its correlation with the economic viability of the dairy industry, early diagnosis of udder pathologies is particularly important. Equally important is the development of effective genetic analysis methods to identify genes potentially associated with udder health in dairy animals. In this regard, several studies have been carried out with a specific focus on genes associated with this aspect of animal health (Tribout et al., 2020, Wu et al., 2015, Brzáková et al., 2021, Jonas et al., 2011). In particular, variants of genes related to immunity have been shown to play a fundamental role in udder health and in modulating host susceptibility to infectious diseases (Khan et al., 2023, Singh et al., 2014, Wang et al., 2018)

Of all the genes studied, *MBL2* is one of the most promising, as the corresponding protein plays a key role in the innate immune system, binding to a range of sugars including N-acetyl-D-glucosamine, mannose, N-acetyl-mannosamine, fucose and glucose.

When mannose-binding lectin (MBL) binds to its ligands, it activates the complement system. However, it can also opsonize bacteria without the activation of the complement system (Thiel et al. 1997, Nepomuceno et al. 1997).

The role of *MBL2* gene polymorphisms in infectious diseases and mastitis in dairy animals is supported by several scientific evidence. Wang et al. (2012) and Zhao et al. (2012) showed that polymorphisms on exon 1 of the *MBL*2 gene in cows are associated with both Somatic Cell Score (SCS) and milk quality parameters.

SNPs and the combination of *MBL2* haplotypes in Italian Mediterranean river buffaloes have been reported to increase susceptibility to *Brucella abortus* (Capparelli et al., 2008). Similarly, certain SNPs and haplotypes were significantly more common in animals with mastitis, as shown by Shergojry et al. (2023). So far, no studies correlating SNPs along the *MBL2* gene and susceptibility to mastitis in sheep have been carried out. In fact, Zhao et al. (2011) only identified four patterns and correlated them with the corresponding protein expression. In particular, serum MBL levels for pattern A (mean 2384 g/L) and pattern B (mean 2309 g/L) were significantly lower ($P < 0.05$) than those for pattern C (mean 2949 g/L) and pattern D (mean 2819 g/L). Serum levels of MBL did not differ significantly between pattern A and pattern B, nor between pattern C and pattern D ($P > 0.05$). Cosenza et al. (2012) only studied the variability within the promoter region and exon 1 of Sardinian sheep. Furthermore, the *MBL2* gene in ruminant species shares common structural features.

Considering these findings, the aim of this study was to characterize for the first time the exon 1 region of *MBL2* and to identify allelic variants and their association with susceptibility to mastitis and milk quality traits in Comisana sheep breed. The Comisana breed originating from the crossbreeding of Maltese and Sicilian breeds in the late 19th and early 20th century, is a native sheep breed of medium to large in size and possesses all the morphological traits of a good dairy breed. It is well-suited to the semi-arid Mediterranean environment, making it a valuable resource particularly in the marginal areas of Southern Italy (Selvaggi et al 2017).

2. Materials and Methods

Animals

Blood samples were collected from 66 lactating Comisana sheep from 2 herds, reared in Lazio region and registered in the herd book. The blood samples were collected with anticoagulant EDTA (K3) and stored at -20°C until extraction. The information on positivity/negativity to the bacteriological test, Somatic Cell Count (SCC) and Differential Somatic Cell Count (DSCC) was provided by the Milk and Diagnostic Laboratory of Istituto Zooprofilattico Sperimentale del Lazio e della Toscana (IZSLT - Experimental Zooprophylactic Institute of Latium and Tuscany) “M.Aleandri”.

On the basis of the bacteriological data, SCC and DSCC, the animals were divided into two groups: negative to mastitis (n = 49) that is negative to bacterial culture, SCC <400,000 and DSCC <60%; and positives to mastitis (n = 18) that is positive to bacterial culture on one or both udders, SCC >400,000 and DSCC >60%.

Genomic DNA extraction, PCR amplification and Sequencing

Genomic DNA was extracted from peripheral blood samples using the Wizard Genomic DNA Purification kit (Promega Corporation) according to the protocol of the manufacturer. To assess the quality and quantity of DNA, the samples were measured using a bio-photometer. Only samples with 260/280 ratio between 1,6 and 1,8 and DNA>30ng/ul have been used.

The primer pair used for amplification were from Zhao et al. (2011) and were designed to amplify a 194bp fragment in the exon 1 region of the *MBL2* gene (Table 1).

The PCR mix contained 1.25U of Taq G2 DNA polymerase (Promega), 200μM for each dNTP (Promega), 100-150ng of DNA template, 0.1 μM of each primer (Invitrogen), with the addition of MilliQ water up to 15μl. The amplification was performed using a BIO-RAD MyCycler thermocycler with the following conditions: pre denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 57°C for 20 s for each set of primer and extension at 72°C for 20 s, and a final extension at 72°C for 10 min. 5 μL of PCR product was electrophoresed on 1.8% agarose gel to check the quality and specificity of amplified DNA product. The amplicons were sequenced using the Sanger method by an external service. The resulting sequences were analyzed with the DNASTAR SeqMan program for characterization and identification of SNPs in the target regions of the gene. The sequence of the amplicons was compared to the reference sequences of the sheep *MBL2* gene with accession numbers FJ977629 and AM933378, but also to that of Zhao et al. (2011).

Gene region	Primer	Sequence (5'- 3')	Fragment length (bp)
Exon 1	MBLF	5′-CGCTGTTTACATCACTTCCT-3′	194bp
	MBLR	5′-CACTGTACCTGGTTCTCCCT-3′	

Table 1: Primers used for amplification.

Statistical Analysis

Genotypic frequencies, allelic frequencies, polymorphism information content (PIC), heterozygosity (He), effective number of alleles (Ne) and Hardy-Weinberg equilibrium (HWE) were analyzed using POPGENE32 (ver.1.31) (Yeh et al., 1999). The data were analyzed using IBM SPSS Statistic software, 29.0.1.0.(171) Version. General linear models (GLM) for repeated measures were used for comparisons genotypes with the milk performance and quality parameters. Bonferroni's tests were used for post hoc multiple comparisons. Data were expressed as means ± standard deviation. Statistical significance was assigned at $p \leq 0.05$. The statistical model for haplotype association with mastitis and milk parameters included the genotype as the dependent variable (six levels), number of lactations, duration of lactation, average of daily milk production, somatic cell score, the random animal effect, and the residual error term as fixed effects. Subsequently, the same animals were grouped based on their genotype at the *MBL2* locus into six different clusters, AA, AC, AD, CC, CD, DD. The model used was the same model described above. The values were considered significant at $p < 0.05$ and presented as the least squares means ± standard errors in both cases. If more than two groups were compared, a Bonferroni test was used for multiple testing.

3. Results and Discussions

Sheep *MBL2* gene sequences available in GenBank (accession number FJ977629.1, AM933378.1) were used as reference for comparison with the sequences obtained from Comisana sheep. In addition, the sequences were compared with those found by Zhao et al. (2011) to see if the haplotypes were the same.

Sequencing revealed a total of 3 haplotypes in the tested population, namely A, C and D according to the nomenclature proposed by Zhao et al. (2011). Moreover, pattern A showed a nucleotide change at position 1080G>A (Fig.1) thus we named it A1. This substitution is responsible for an amino acid change from arginine to glutamine at p.28 of the amino acid sequence (Fig. 1). As arginine is a basic amino acid, whereas glutamine is polar, this could lead to a change in the structure of the resulting protein, which in turn could lead to a change in its activity.

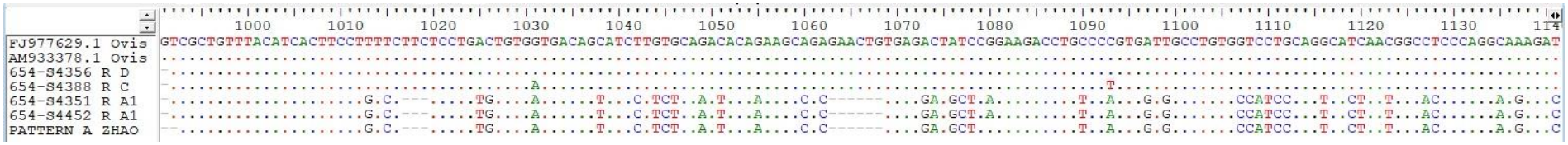

Fig. 1: Comparison of the results of sequencing with reference sequences.

In the Comisana sheep, A1 had a frequency of 0.416 and is the most common in the population analyzed, followed by D (0.371) and C (0.212). B was not observed (Table 2).

In the Hu sheep, pattern A was the most frequent (0.457), followed by pattern B (0.267), C (0.162), and finally D (0.114) (Zhao et al., 2011). In both breeds, the A variant was the most abundant, suggesting that it is probably the most advantageous variant in terms of resistance to infectious diseases.

Furthermore, the absence of the B variant in the breed studied and its presence as the second most frequent allele in the Hu sheep could suggest an origin and evolution specific to Asian populations. This hypothesis should be validated with more specific analyses like phylogenetic ones.

	Genotype						**Allele Frequency**			**Population Indices**		
	A1A1	A1C	CC	A1D	CD	DD	A1	C	D	Ho	He	Fis
Obs	15 (0,227)	21 (0,318)	2 (0,031)	4 (0,061)	3 (0,045)	21 (0,318)	0,416	0,212	0,371	0,4242	0,6485	0,3408
Exp	11,34	11,75	2,9	20,57	10,47	8,98						
χ^2=43,5												
p=0												
d.f.=5												

Table 2: Genotype, allele frequencies and population indices of the Comisana sheep.

The data shown in Table 2 highlight that the genotype distribution in Comisana sheep deviate from the Hardy-Weinberg equilibrium (χ^2=43,5). The Ho, He and Fis parameters, which indicate the prevalence of homozygosity at the expense of heterozygosity, provide further confirmation.

The population analysis on the two separate groups (healthy *vs.* mastitic) confirmed the deviation from Hardy-Weinberg equilibrium (χ^2=32.169 healthy group, Table 3) (χ^2=18.14 mastitic group, Table 4). Furthermore, it is observed that the frequency of the A1 allele remains similar in both groups, while the C and D alleles are distributed differently: the former is more represented in the healthy, the latter in the sick. This suggests a protective effect of the C allele against mastitis, although the effect is not statistically significant.

	Genotype						Allele Frequency			Population Indices		
	A1A1	A1C	CC	A1D	CD	DD	A1	C	D	Ho	He	Fis
Obs	8 (0,163)	21 (0,428)	2 (0,041)	4 (0,082)	2 (0,041)	12 (0,245)	0,418	0,275	0,306	0,551	0,662	0,159
Exp	8,45	11,41	3,62	12,68	8,35	4,48						
χ^2=32,169												
p=0,00												
d.f.=5												

Table 3: Genotype, allele frequencies and population indices in healthy animals.

	Genotype						Allele Frequency			Population Indices		
	A1A1	A1C	CC	A1D	CD	DD	A1	C	D	Ho	He	Fis
Obs	7 (0,412)	0	0	0	1 (0,058)	9 (0,529)	0,412	0,029	0,558	0,058	0,533	0,886
Exp	2,76	0,42	0	8,06	0,58	5,18						
χ^2=18,14												
p=0,0004												
d.f.=5												

Table 4: Genotype, allele frequencies and population indices in mastitic animals.

Genotype	**Group**		**Tot**
	Healthy	**Mastitic**	
A1A1	8	7	15
A1C	21	0	21
A1D	4	0	4
CC	2	0	2
CD	2	1	3
DD	12	9	21
Tot	49	17	66

Table 5: Distribution of genotypes in healthy and mastitic group.

It was also found that the A1C genotype is completely absent in the mastitic group, while it is the most represented in the healthy one, as shown in the Table 5.

Furthermore, the allele frequencies showed a higher presence of the D allele (0.558) in the mastitic group. This would contradict Zhao et al. (2011), who show that the amount of the corresponding serum protein was significantly higher in subjects with the C and D pattern, making them likely less susceptible to infectious diseases. It should be emphasized that Zhao et al. (2011) based his

conclusions on what observed in humans and did not carry out studies on the functionality of the protein or even a correlation with disease/health status in animals, so it is not possible to say with certainty that the patterns responsible for higher protein production protect against the disease risk. Given the difference in nucleotide and therefore amino acid sequence, the protein from pattern A1 may be more active than the protein produced by patterns C and D, which could explain the high frequency of pattern D in the mastitic group.

Finally, it should be noted that pattern A1 was found for the first time in this work and it cannot be excluded that its association with C pattern in the genotype confers a greater resistance than the A variant found by Zhao et al. (2011).

Genotype	SCC	DSCC
A1A1	1,673.33 ± 2,907.60[b]	51,66 ± 38,88
A1C	73,76 ± 40,44[a]	34,63 ± 29,92
DD	1,956.61 ± 4,148.37[b]	52,03 ± 32,63

Table 6: SCC and DSCC of Comisana sheep milk according to the genotype. a,b $p<0,05$.

This hypothesis could explain the statistical significance observed in the association analysis between A1 genotype and a low SCC (Table 6). This would be consistent with what has been found in cows (Wang et al., 2012, Zhao et al., 2012) i.e. that variants along exon 1 of the *MBL2* gene are associated with SCS and milk quality parameters.

4. Conclusions

In this work, the sequence of exon 1 of the *MBL2* gene was for the first time analyzed in Comisana sheep and a new variant (A1) with effect on the coded protein was observed.

Statistical analyses revealed interesting differences between the genotypes found in healthy and diseased animals. Furthermore, the association of genotype A1 with SCC parameter in the diseased animals suggests a relationship between the variant found and the mastitis indicators.

Therefore, it would be appropriate to investigate the role of variants of the *MBL2* gene in sheep regarding resistance to the infectious diseases, by increasing the number of animals studied and by considering other sheep breeds.

If confirmed, these data would provide sheep breeders with an additional tool useful to make more informed decisions about the management of their flocks: favoring more resistant animals by adopting targeted mating strategies that increase the frequencies of positive genotype and detecting animals with susceptible genotype and monitoring them more closely.

In summary, increased awareness and knowledge of the genetic basis of animal health issues can lead to significant improvements in health, welfare and productivity, thereby contributing to breeding success.

Funding

This study is part of the Current research project 2020 entitled "Use of markers for the early detection of subclinical mastitis, study of genes related to genetic resistance to mastitis and control of somatic cells in milk in sheep and buffalo herds" - Istituto Zooprofilattico Sperimentale del Lazio e della Toscana "M. Aleandri".

References

1. Bhosale,R.R.; Osmani,R.A.; Ghodake, P.P.; Shaikh, S.M.; Chavan, S.R. Review Article Mastitis: An Intensive Crisis in Veterinary Science. 2014. Int. J. Pharma Res. Health Sci., 2(2): 96–103.
2. Brzáková, M.; Rychtářová, J.; Čítek, J.; Sztankóová, Z. A Candidate Gene Association Study for Economically Important Traits in Czech Dairy Goat Breeds. *Animals* **2021**, *11* (6), 1796.https://doi.org/10.3390/ani11061796.
3. Capparelli, R.; Parlato, M.; Amoroso, M. G.; Roperto, S.; Marabelli, R.; Roperto, F.; Iannelli, D. Mannose-Binding Lectin Haplotypes Influence Brucella Abortus Infection in the Water Buffalo (Bubalus Bubalis). Immunogenetics 2008, 60 (3–4), 157–165. https://doi.org/10.1007/s00251-008-0284-4.
4. Cosenza, G.; Pauciullo, A.; Mancusi, A.; D'Avino, A.; Colimoro, L.; Di Berardino, D.; Ramunno, L. Identification of Single Nucleotide Polymorphisms in the Nicastrese Goat and Sardinia Sheep Mannose-Binding Lectin. Biochem Genet 2012, 50 (1–2), 73–83.https://doi.org/10.1007/s10528-011-9461-5.
5. Halasa, T.; Huijps, K.; Østerås, O.; Hogeveen, H. Economic Effects of Bovine Mastitis and Mastitis Management: A Review. Veterinary Quarterly 2007, 29 (1), 18–31. https://doi.org/10.1080/01652176.2007.9695224.
6. Jonas, E.; Thomson, P. C.; Hall, E. J.; McGill, D.; Lam, M. K.; Raadsma, H. W. Mapping Quantitative Trait Loci (QTL) in Sheep. IV. Analysis of Lactation Persistency and Extended Lactation Traits in Sheep. *Genet Sel Evol* **2011**, *43* (1), 22. https://doi.org/10.1186/1297-9686-43-22.
7. Khan, M. Z.; Wang, J.; Ma, Y.; Chen, T.; Ma, M.; Ullah, Q.; Khan, I. M.; Khan, A.; Cao, Z.; Liu, S. Genetic Polymorphisms in Immune- and Inflammation-Associated Genes and Their Association with Bovine Mastitis Resistance/Susceptibility. Front. Immunol. 2023, 14, 1082144. https://doi.org/10.3389/fimmu.2023.1082144.
8. Libera, K.; Konieczny, K.; Grabska, J.; Smulski, S.; Szczerbal, I.; Szumacher-Strabel, M.; Pomorska-Mól, M. Potential Novel Biomarkers for Mastitis Diagnosis in Sheep. Animals 2021, 11 (10), 2783. https://doi.org/10.3390/ani11102783.
9. Martí-De Olives, A.; Peris, C.; Molina, M. P. Effect of Subclinical Mastitis on the Yield and Cheese-Making Properties of Ewe's Milk. Small Ruminant Research 2020, 184, 106044. https://doi.org/10.1016/j.smallrumres.2019.106044.
10. Nepomuceno, R. R.; Henschen-Edman, A. H.; Burgess, W. H.; Tenner, A. J. cDNA Cloning and Primary Structure Analysis of C1qRP, the Human C1q/MBL/SPA Receptor That Mediates Enhanced Phagocytosis In Vitro. Immunity 1997, 6 (2), 119–129. https://doi.org/10.1016/S1074-7613(00)80419-7.
11. Petrovski, K. R.; Trajcev, M.; Buneski, G. A Review of the Factors Affecting the Costs of Bovine Mastitis : Review Article. J. S. Afr. Vet. Assoc. 2006, 77 (2), 52–60. https://doi.org/10.4102/jsava.v77i2.344.
12. Selvaggi, M.; D'Alessandro, A. G.; Dario, C. Environmental and Genetic Factors Affecting Milk Yield and Quality in Three Italian Sheep Breeds. Journal of Dairy Research 2017, 84 (1), 27–31.https://doi.org/10.1017/S0022029916000765.
13. Shergojry, S. A.; Verma, A.; Ghani, M.; Gupta, I. D.; Mir, N. A. Identification of Genetic Polymorphism of the MBL2 Gene and Its Association with Clinical Mastitis in Murrah Buffaloes. J Genet 2023, 102 (1), 21.https://doi.org/10.1007/s12041-023-01419-9.
14. Singh, U.; Deb, R.; Alyethodi, R. R.; Alex, R.; Kumar, S.; Chakraborty, S.; Dhama, K.; Sharma, A. Molecular Markers and Their Applications in Cattle Genetic Research: A Review. Biomarkers and Genomic Medicine 2014, 6 (2), 49–58. https://doi.org/10.1016/j.bgm.2014.03.001

15. Thiel, S.; Vorup-Jensen, T.; Stover, C. M.; Schwaeble, W.; Laursen, S. B.; Poulsen, K.; Willis, A. C.; Eggleton, P.; Hansen, S.; Holmskov, U.; Reid, K. B. M.; Jensenius, J. C. A Second Serine Protease Associated with Mannan-Binding Lectin That Activates Complement. Nature 1997, 386 (6624), 506–510. https://doi.org/10.1038/386506a0.
16. Tribout, T.; Croiseau, P.; Lefebvre, R.; Barbat, A.; Boussaha, M.; Fritz, S.; Boichard, D.; Hoze, C.; Sanchez, M.-P. Confirmed Effects of Candidate Variants for Milk Production, Udder Health, and Udder Morphology in Dairy Cattle. *Genet Sel Evol* **2020**, *52* (1), 55. https://doi.org/10.1186/s12711-020-00575-1.
17. Wang, M.; Song, H.; Zhu, X.; Xing, S.; Zhang, M.; Zhang, H.; Wang, X.; Yang, Z.; Ding, X.; Karrow, N. A.; König, S.; Mao, Y. *Toll-like Receptor* 4 Gene Polymorphisms Influence Milk Production Traits in Chinese Holstein Cows. *Journal of Dairy Research* **2018**, *85* (4), 407–411.https://doi.org/10.1017/S0022029918000535.
18. Wang, X.; Ju, Z.; Huang, J.; Hou, M.; Zhou, L.; Qi, C.; Zhang, Y.; Gao, Q.; Pan, Q.; Li, G.; Zhong, J.; Wang, C. The Relationship between the Variants of the Bovine MBL2 Gene and Milk Production Traits, Mastitis, Serum MBL-C Levels and Complement Activity. Veterinary Immunology and Immunopathology 2012, 148 (3–4), 311–319. https://doi.org/10.1016/j.vetimm.2012.06.017.
19. Wu, X.; Lund, M. S.; Sahana, G.; Guldbrandtsen, B.; Sun, D.; Zhang, Q.; Su, G. Association Analysis for Udder Health Based on SNP-Panel and Sequence Data in Danish Holsteins. *Genet Sel Evol* **2015**, *47* (1), 50. https://doi.org/10.1186/s12711-015-0129-1.
20. Yeh, F.C., Yang, R.C. and Boyle, T. (1999) POPGENE Version 1.32: Microsoft Window-Based Freeware for Population Genetics Analysis. University of Alberta, Edmonton.
21. Zhao, F.; Zhao, Z.; Yan, G.; Wang, D.; Ban, Q.; Yu, P.; Zhang, W.; Luo, Y. Polymorphisms in Mannose-Binding Lectin (MBL) Gene and Their Association with MBL Protein Levels in Serum in the Hu Sheep. Veterinary Immunology and Immunopathology 2011, 140 (3–4), 297–302.https://doi.org/10.1016/j.vetimm.2010.12.009.
22. Zhao, Z. L.; Wang, C. F.; Li, Q. L.; Ju, Z. H.; Huang, J. M.; Li, J. B.; Zhong, J. F.; Zhang, J. B. Novel SNPs of the Mannan-Binding Lectin 2 Gene and Their Association with Production Traits in Chinese Holsteins. Genet. Mol. Res. 2012, 11 (4), 3744–3754.https://doi.org/10.4238/2012.October.15.6.

CONCLUSIONS

The results of this study were particularly interesting, as all the genes studied were shown to influence milk quality characteristics, although not all of them were directly involved in the synthesis of its constituents, as *MBL2* gene. In fact, the *MBL2* gene was identified as a key factor in disease resistance, influencing animal welfare and thus indirectly milk quality, as it was related to the SCC (Somatic Cell Count) parameter.

The data collected not only provide information on the genetic characteristics of breeds and species that contribute to the identity of an entire population, but also provide a solid scientific basis for a more conscious approach to species and/or breed conservation.

Knowledge of the variants of the genes that have a direct (*CSN1S1*, *CSN2*, *DGAT1*) and indirect (*MBL2*) effect on production, as well as their distribution in the populations of interest, is essential for optimizing management practices and developing appropriate plans for the valorization of their products, and thus for protecting the market position of these products in order to guarantee the economic sustainability of their breeding and thus their survival, helping to reverse the demographic and socio-economic decline of rural areas.

The aim is essentially to preserve animal biodiversity, which now more than ever represents a fundamental genetic reservoir for facing current and future challenges in the livestock and environmental sectors.

In conclusion, knowledge of the genetic and phenotypic characteristics of indigenous breeds is fundamental for the development of appropriate conservation and management plans, ensuring the sustainability of livestock production in the long term.

OTHER WORKS

genes

MDPI

Case Report

Demonstration of Parthenogenetic Reproduction in a Pet Ball Python (*Python regius*) through Analysis of Early-Stage Embryos

Francesco Di Ianni [1,†], Sara Albarella [2,†], Alessandro Vetere [1,*], Marco Torcello [3], Michela Ablondi [1], Mariagiulia Pugliano [2], Susanna Di Mauro [3], Pietro Parma [4] and Francesca Ciotola [2]

[1] Department of Veterinary Science, Strada del Taglio 10, 43121 Parma, Italy
[2] Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Via Delpino 1, 80137 Naples, Italy
[3] Ambulatorio Veterinario Dott. Di Mauro, Via Parini 8, 24043 Caravaggio, Italy
[4] Department of Agricultural and Environmental Sciences, Via Celoria 2, 20133 Milano, Italy
* Correspondence: alessandro.vetere@unipr.it
† These authors contributed equally to this work.

Abstract: Parthenogenesis is an asexual form of reproduction, normally present in various animal and plant species, in which an embryo is generated from a single gamete. Currently, there are some species for which parthenogenesis is supposed but not confirmed, and the mechanisms that activate it are not well understood. A 10-year-old, wild-caught female ball python (*Python regius*) laid four eggs without any prior contact with a male. The eggs were not incubated and, after 3 days, were submitted to the University of Parma for analysis due to the suspicion of potential embryo presence. Examination of the egg content revealed residual blood vessels and a small red spot, indicative of an early-stage embryo. DNA was extracted from the three deceased embryos and from the mother's blood, five microsatellites were analyzed to ascertain the origin of the embryos. The captive history data, together with the genetic microsatellite analysis approach, demonstrated the parthenogenetic origin of all three embryos. The embryos were homozygous for each of the maternal microsatellites, suggesting a terminal fusion automixis mode of development.

Keywords: parthenogenesis; *P. regius*; VNTR; PCR; genetic; reptiles

Citation: Di Ianni, F.; Albarella, S.; Vetere, A.; Torcello, M.; Ablondi, M.; Pugliano, M.; Di Mauro, S.; Parma, P.; Ciotola, F. Demonstration of Parthenogenetic Reproduction in a Pet Ball Python (*Python regius*) through Analysis of Early-Stage Embryos. *Genes* **2023**, *14*, 1744. https://doi.org/10.3390/genes14091744

Academic Editor: Tingxian Deng

Received: 27 June 2023
Revised: 19 August 2023
Accepted: 29 August 2023
Published: 31 August 2023

1. Introduction

Parthenogenesis (from the two Greek words *parthenos*, meaning virgin, and *genesis*, meaning origin) has been defined as a form of 'asexual' reproduction in which a female reproduces without the participation of a male [1–5]. The term 'parthenogenesis' was later defined by Richard Owen in 1849 as "procreation without the immediate influence of the male". In fact, the embryo forms without male fertilization by sperm [6–8]. These embryos do not always develop into new individuals because they often suffer from high mortality; therefore, the definition of parthenogenesis does not consider the development of new individuals [1] but only of embryos [9]. Parthenogenetic events have been observed in both plants and animals [4,5]. Several mechanisms of parthenogenesis have been demonstrated. Meiosis occurs in automictic or meiotic parthenogenesis, but the mother's chromosomal constitution is reestablished through various mechanisms. Some of these mechanisms result in homozygosity at all loci, while others transmit the mother's genome intact to the offspring. On the other hand, in apomictic or mitotic parthenogenesis, eggs are formed through a series of mitotic divisions, effectively bypassing the process of meiosis [10]. Unlike asexual reproduction, in which new individuals are formed from somatic cells, parthenogenesis involves the development of egg cells with meiotic recombination of the genetic material coming from parents as an "incomplete form of sexual reproduction" [3,9].

In most cases, as there is no genetic recombination, the progeny is genetically identical to its virgin mother [11]. Since parthenogenesis can occur in different ways within the life cycles of different animals, three classifications have been proposed according to different criteria, namely reproduction, sex determination and cytological mechanisms [1,9]. In a broad sense, parthenogenetic lineages are typically classified as either 'generalists' or 'specialists', hinging on their possession of a broad ecological niche and adaptability to diverse environments, or their exceptional fitness within a constrained set of surroundings [5]. Parthenogenesis can be divided into (i) tychoparthenogenesis, when unfertilized eggs occasionally develop through parthenogenesis; (ii) facultative parthenogenesis, when the eggs may be fertilized or develop parthenogenetically and (iii) obligate parthenogenesis, when eggs always develop parthenogenetically. Among squamate reptiles, approximately 30 unisexual species have been formally named as belonging to distinct families, including Typhlopidae, Gekkonidae, Gymnophthalmidae, Xantusiidae, Teiidae, Scincidae and Lacertidae [12,13]. In recent years, studies focused on facultative parthenogenesis in snakes have produced innovative and field-advancing outcomes, substantially enriching our knowledge and comprehension of this fascinating reproductive phenomenon in these reptiles. While obligate parthenogesis is reported in the brahminy blind snake (*Indotyphlops braminus*) through the mechanism of premeiotic endoreplication, facultative parthenogenesis has been shown in other snakes and can be attributed to terminal fusion automixis. Moreover, it has been observed throughout snake evolution, starting with ancient boas (Boidae), pythons (Pythonidae) and Caenophidia [11,12]. This form of reproduction has been detected in multiple lineages of more modern or 'advanced' snakes [11,13–15] with several genetically confirmed or anecdotical cases [15–23]. Other species involved in similar studies were the western terrestrial garter snake (*Thamnophis elegans*), checkered garter snake (*T. marcianus*), timber rattlesnake (*Crotalus horridus*), Aruba island rattlesnake (*C. unicolour*), arafura file snake (*Acrochordus arafurae*) [24,25], reticulate python (*Malayopython reticulatus*), ball python (*P. regius*) [11] and Burmese python (*Python molurus bivittatus*) [26]. The last-mentioned case was the first one reported in the Pythonidae family, although a more detailed analysis has highlighted some errors of interpretation [15,22]. In another report, a 22-year-old captive Brazilian rainbow boa (*Epicrates cenchria cenchria*) gave birth to four offspring after being housed with a vasectomized male. The male, female and four offspring were genotyped for short tandem repeat (STR) DNA markers. None of the offspring exhibited the specific STR allele found in the possible sire. Furthermore, all the offspring displayed homozygosity at each assessed STR locus, providing evidence for parthenogenetic reproduction [20]. Booth W et al. provided the initial scientifically established evidence of facultative parthenogenesis in a primitive snake from the Boidae family (*Boa constrictor*). This discovery was made possible through the application of microsatellite DNA fingerprinting. The increased homozygosity observed in the offspring compared to the mother implies that the mechanism behind this phenomenon might be terminal fusion automixis. Notably, no male offspring were generated, potentially pointing towards maternal sex chromosome hemizygosity [22]. Whether parthenogens demonstrate reproductive competency is pivotal in understanding the evolutionary and ecological importance of facultative parthenogenesis in wild squamates and other vertebrates. Extensive research on domestic turkeys has recorded the reproductive capability of parthenogens [6,10]. Also, Straube et al. documented reproductive competency in a parthenogenetic whitespotted bamboo shark, where the parthenogen itself produced a parthenogen [12]. In that study, a captive female produced multiple parthenogens. Unexpectedly, out of a collective of nine parthenogens, a solitary specimen exhibited external claspers typically linked to the male gender among chondrichthyans. Upon closer examination through dissection, this particular specimen showed misshapen or missing internal sexual organs. Notably, the existence of claspers in this study raises questions about the previously assumed mechanism of sex determination in this species of shark. In this study, a possible case of parthenogenesis in a ball python (*P. regius*) was analyzed by genotyping five specific microsatellites. The genetic comparisons between

mother and embryos demonstrated the parthenogenetic origin of the latter, showing that this reproductive strategy is not rare in this species [11].

2. Material and Methods

2.1. Animals

A 10-year-old, wild-caught, 1.3 kg female ball python (*P. regius*) laid 4 eggs without ever being in contact with a male. The eggs were promptly removed by the owner and not incubated. After 3 days, an expert veterinarian was asked for a complete check-up of the snake. The snake was born in captivity and acquired at approximately 3 months old; husbandry and diet were considered appropriate for the species [27]. The veterinarian analyzed the egg content, finding residual blood vessels and a small red spot compatible with an early-stage embryo. All the eggs were then submitted to the University of Parma for analysis. Blood (1.3 mL) was collected from the female python by jugular venipuncture and equally divided into two test tubes, one with lithium heparin and one with kEDTA. Genetic analyses were conducted on three embryos, labeled 3068, 3069 and 3144.

2.2. DNA Extraction and Amplification

DNA was extracted from the three embryos and from the blood of the mother using the commercial Wizard Genomic DNA extraction kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. DNA yield and purity were measured on a BioPhotometer (Eppendorf, Hauppauge, NY, USA) and showed the following requirements: yield > 30 ng/µL, 260/280 ratio > 1.7 and 260/230 ratio > 1.8. Before the PCR amplification, all the genomic DNA samples were diluted at 20 ng/µL in nuclease-free water and immediately used. Each amplification was carried out in a 25 µL final volume containing 100 ng of genomic DNA, a 1× PCR buffer, 200 µM of each dNTP, 10 pmol of each primer, and 1 U of GoTaq® G2 Flexi DNA Polymerase (Promega–Madison, Fitchburg, WI, USA).

The thermal conditions were as follows: 98 °C for 30 s, 30 cycles at 98 °C for 10 s, annealing at 64 °C for 30 s and extension at 72 °C for 30 s. A final extension was carried out at 72 °C for 10 min. The amplification products were later verified by electrophoresis on a 2% agarose gel (Bio-Rad, Hercules, CA, USA) in a 0.5X TBE buffer and stained with SYBR®green (Lonza Rockland, Inc., Rockland, ME, USA).

2.3. Choice of Microsatellites

To verify the genomic identity between mother and embryos, microsatellites previously used in *Python regius* were chosen (Table 1). The amplification products were then purified with ExoSap and sequenced with a Brilliant Dye Terminator 1.1 kit (Applied Biosystems™, Waltham, MA, USA) and a 3730xl DNA Analyzer (Applied Biosystems™, Waltham, MA, USA). In this way, it was possible to count the number of repetitions of the repeated sequences and assign the correct allele.

Table 1. Characteristics of the microsatellites used.

Name	All [a]	Ho [b]	He [c]	PIC [d]	Primers	Reference
MS9	7	0.78	0.79	0.75	5′-CAGTGGGCTTGAGATTGAC-3′ 5′-CCATTCCTTAAAACACTCTCACTC-3′	[27]
MS5	25	0.91	0.93	0.92	5′-TAGGGTGTCAGTCATTGCTC-3′ 5′-TGGCATCCAGCAGTCATAG-3′	[27]
MS16	10	0.75	0.82	0.78	5′-GAGTTCTGGTCTTGCTTTCG-3′ 5′-CAGGTACAACTTTCTCCAAC-3′	[27]
Pmbl-12	4	0.55	0.59	n.a. [e]	5′-GCCACGTCTAAGGTTGAGC-3′ 5′-AAAGCAGGTCTCTGTTGGG-3′	[28]
KE955519	8	0.75	0.72	0.68	5′-ATTTTAGCTGCAGGCTGTGG-3′ 5′-TCTGCTAGGGCAAAACTGGG-3′	[27]

[a] number of alleles; [b] observed heterozygosity; [c] expected heterozygosity; [d] polymorphism information content; [e] not available.

2.4. Statistical Analysis

To statistically support the parthenogenetic event, we calculated the probability of sexual reproduction in the presence of sperm storage. The calculation procedure was carried out according to Booth et al. (2014) [11].

3. Results

The genetic material obtained made it possible to genotype all three embryos for the five microsatellites considered. The results obtained are shown in Table 2. For the 15 microsatellite–embryo combinations, homozygosity was observed at all loci.

Table 2. Genotypes of the mother and offspring at five microsatellite loci.

	MS9	MS5	MS16	Pmbl-12	KE9555519
Mother	164/168	346/362	350/366	480/480	346/348
3068	164/164	362/362	366/366	480/480	346/346
3069	168/168	346/346	366/366	480/480	348/348
3144	164/164	362/362	366/366	480/480	346/346

Considering that the allele frequencies for the selected microsatellites in the population of origin of the studied female are not known, it was not possible to calculate the probability of the parthenogenetic event according to the classical scheme. In this case, we calculated the probability of sexual reproduction in the presence of long-term sperm storage, as reported by Booth et al. (2014) [11]. Results are reported in Table 3.

Table 3. Probability of sexual reproduction.

	Number of Offspring	Number of Maternally Homozygous Loci	Number of Maternally Heterozygous Loci	Probability of Long-Term Sperm Storage (per Individual)	Probability of Long-Term Sperm Storage (per Clunch)
Mother	3	1	4	0.001953	7.45×10^{-9}

4. Discussion

This study can be considered the fourth report of facultative parthenogenesis (FP) in a captive-bred female ball python (*P. regius*), as previously reported by Booth et al. [11,15]. Parthenogenesis is also documented in oviparous snakes [26], while obligate parthenogenesis is present in only a single lineage of an extant basal scolecophidian snake, the common blind snake (*I. braminus*) [29]. Currently, more than a million ball pythons (*P. regius*) are being kept in captivity and zoos, making them one of the most favored choices among pet snake enthusiasts [30]. Indigenous to Central and Western Africa, these snakes have a typical length range of 0.9–1.8 m (3–6 feet) and can thrive for over 30 years in captivity [30]. Their generally calm temperament enhances their appeal as a sought-after pet choice. Since 1976, Ghana and Togo have been the primary sources, contributing nearly 100% of the specimens that are predominantly exported to the USA and EU. Consequently, ball pythons have emerged as the most extensively traded CITES-listed species originating from Africa, with hundreds of thousands of these creatures being involved in international trade annually [31]. In their natural habitat, royal pythons primarily engage in breeding activities spanning from mid-September to mid-November, aligning with the duration of the rainy season. These pythons follow an oviparous reproductive strategy, with females typically laying an average clutch of six eggs. However, the range of clutch sizes can vary from as few as one egg to as many as eleven [31]. Currently, there is only one study regarding FP in ball pythons [11]. In this study by Booth et al., complete genomic DNA was isolated from the molted skins of living individuals or muscle tissue obtained from deceased ball python embryos. These samples were sourced from three separate and unrelated egg clutches that were generated within three distinct private collections [11]. After mating, some female snakes can store male sperm within their reproductive tracts for extended

periods of time, laying fertile eggs after years of reproductive inactivity [32–36]. In a study by Booth et al., a rattlesnake (*Crotalus adamanteus)* was found to hold one of the longest genetically verified instances of long-term sperm storage (LTSS) among vertebrates. In that study, it was captured when it was a young adult. Approximately 60 months later, it gave birth to a substantial and healthy litter consisting of 19 offspring, including 9 males and 10 females [23]. In a more recent study by Levine et al., a female western diamond-backed rattlesnake (*Crotalus atrox*) collected from the wild and kept in isolation since her capture in September 1999 gave birth to two healthy litters, with the intervals between the births being approximately one and six years following her capture. An analysis of the genetic makeup of the 2005 litter revealed the presence of paternal genetic contributions in all offspring [35]. As the female ball python under study was kept without a male since it was sexually immature, this mechanism of internal fertilization was excluded. In the study by Groot et al. [26], a histopathologic examination was performed on seven 24-day-old, 18 cm long embryos for sex determination; based on the presence of ovaries, all offspring were identified as female [15]. Considering that pythons possess sex chromosomes in the form of XX/XY, where XX corresponds to females, the embryos in this case, although of insufficient size for sex determination, would necessarily have been of the female sex [37]. In this study, we analyzed three 2 mm long early-stage snake embryos. The results of this study are in line with those of Booth et al. [15], where all embryos showed homozygosity for the tested microsatellites, despite some being heterozygous in the mother. The presence of only one subset of maternal microsatellites might be interpreted as a sign of terminal fusion automixis, considered to be the developmental mechanism mainly underlying vertebrate FP and characterized by the restoration of diploidy through the fusion or duplication of meiotic products. Among the automictic modes, terminal fusion is the most common mechanism [1,12,35,36], under which the egg nucleus fuses with its second polar body, resulting in highly reduced levels of heterozygosity throughout the genome.

5. Conclusions

The case report described in this paper, relating to the microsatellite analyses carried out in a female ball python and in three early-stage embryos, demonstrates with certainty a phenomenon of FP in their origin. As of today, the use of few-day embryos as a source of genetic material to demonstrate parthenogenesis has not been previously documented, presenting new opportunities to utilize this biological material for this purpose. Furthermore, the increasingly frequent identification of cases of FP in *P. regius* demonstrates that this species represents an excellent natural animal model for studying the mechanisms underlying FP.

Author Contributions: F.D.I. and S.A. contributed equally and share the first authorship. Literature review: A.V., F.C., M.T., P.P. and S.D.M. Acquisition of data: F.C., P.P., M.A. and M.P. Analysis and interpretation of data: F.C., P.P. and M.A. Drafting the article: A.V., S.A., M.A., P.P. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical approval for the study was given by the University of Parma (Organismo Preposto Al Benessere Degli Animali Comitato Etico Per La Sperimentazione Animale) (PROT. N. 02/CESA/2023).

Informed Consent Statement: Owners gave written consent for their pets' personal or clinical details, along with any identifying images, to be published in this study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no competing interest.

References

1. Lynch, M. Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *Q. Rev. Biol.* **1984**, *59*, 257–290. [CrossRef]
2. Dedukh, D.; Majtánová, Z.; Marta, A.; Pšenička, M.; Kotusz, J.; Klíma, J.; Juchno, D.; Boron, A.; Janko, K. Parthenogenesis as a solution to hybrid sterility: The mechanistic basis of meiotic distortions in clonal and sterile hybrids. *Genetics* **2020**, *215*, 975–987. [CrossRef] [PubMed]
3. Allen, L.; Sanders, K.L.; Thomson, V.A. Molecular evidence for the first records of facultative parthenogenesis in elapid snakes. *R. Soc. Open Sci.* **2018**, *5*, 171901. [CrossRef] [PubMed]
4. Janko, K.; Pačes, J.; Wilkinson-Herbots, H.; Costa, R.J.; Roslein, J.; Drozd, P.; Iakovenko, N.; Rídl, J.; Hroudová, M.; Kočí, J.; et al. Hybrid asexuality as a primary postzygotic barrier between nascent species: On the interconnection between asexuality, hybridization and speciation. *Mol. Ecol.* **2018**, *27*, 248–263. [CrossRef] [PubMed]
5. Koivisto, R.K.K.; Braig, H.R. Microorganisms and parthenogenesis. *Biol. J. Linn. Soc.* **2003**, *79*, 43–58. [CrossRef]
6. Simon, J.C.; Delmotte, F.; Rispe, C.; Crease, T. Phylogenetic relationships between parthenogens and their sexual relatives: The possible routes to parthenogenesis in animals. *Biol. J. Linn. Soc.* **2003**, *79*, 151–163. [CrossRef]
7. Gasanov, E.; Katz, A. Facultative parthenogenesis in a diamondback water snake (*Nerodia rhombifer*). *Russ. J. Herpetol.* **2020**, *27*, 341–347. [CrossRef]
8. Kearney, M. Hybridization, glaciation and geographical parthenogenesis. *Trends Ecol. Evol.* **2005**, *20*, 495–502. [CrossRef]
9. Mittwoch, U. Parthenogenesis. *J. Med. Genet.* **1978**, *15*, 165–181. [CrossRef]
10. Stenberg, P.; Saura, A. Cytology of asexual animals. In *Lost Sex: The Evolutionary Biology of Parthenogenesis*; Schön, I., Martens, K., Dijk, P., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 2009; pp. 63–74.
11. Booth, W.; Schuett, G.W.; Ridgway, A.; Buxton, D.W.; Castoe, T.A.; Bastone, G.; Bennett, C.; McMahan, W. New insights on facultative parthenogenesis in pythons. *Biol. J. Linn. Soc.* **2014**, *112*, 461–468. [CrossRef]
12. Card, D.C.; Vonk, F.J.; Smalbrugge, S.; Casewell, N.R.; Wüster, W.; Castoe, T.A.; Schuett, G.W.; Booth, W. Genome-wide data implicate terminal fusion automixis in king cobra facultative parthenogenesis. *Sci. Rep.* **2021**, *11*, 7271. [CrossRef] [PubMed]
13. Booth, W.; Million, L.; Reynolds, R.G.; Burghardt, G.M.; Vargo, E.L.; Schal, C.; Tzika, A.C.; Schuett, G.W. Consecutive virgin births in the New World boid snake, the *Colombian rainbow boa*, *Epicrates maurus*. *J. Hered.* **2011**, *102*, 759–763. [CrossRef]
14. White, M.J.D. *Animal Cytology & Evolution*; Cambridge University Press: Cambridge, UK, 1954.
15. Booth, W.; Schuett, G.W. The emerging phylogenetic pattern of parthenogenesis in snakes. *Biol. J. Linn. Soc.* **2016**, *118*, 172–186. [CrossRef]
16. Jordan, M.A.; Perrine-Ripplinger, N.; Carter, E.T. An independent observation of facultative parthenogenesis in the copperhead (*Agkistrodon contortrix*). *J. Herpetol.* **2015**, *49*, 118–121. [CrossRef]
17. Vaughan, M.S.; Steele, R.A. Bothrops asper (Terciopelo). Parthenogenetic reproduction. Natural History Notes [Abstract]. *Herpetol. Rev.* **2014**, *45*, 705.
18. Lara-Resendiz, R.A.; Barrios, B.L.; Vega-Perez, A.H.; Centenero-Alcala, E. Natural History Notes: *Boa constrictor* (*Boa constrictor*). Reproduction/Facultative parthenogenesis. *Herpetol. Rev.* **2013**, *44*, 151–152.
19. Reynolds, R.G.; Booth, W.; Schuett, G.W.; Fitzpatrick, B.M.; Burghardt, G.M. Successive virgin births of viable male progeny in the checkered gartersnake, *Thamnophis marcianus*. *Biol. J. Linn. Soc.* **2012**, *107*, 566–572. [CrossRef]
20. Kinney, M.E.; Wack, R.F.; Grahn, R.A.; Lyons, L. Parthenogenesis in a Brazilian Rainbow Boa (*Epicrates cenchria cenchria*). *Zoo Biol.* **2013**, *32*, 172–176. [CrossRef] [PubMed]
21. Booth, W.; Smith, C.F.; Eskridge, P.H.; Hoss, S.K.; Mendelson, I.I.I.J.R.; Schuett, G.W. Facultative parthenogenesis discovered in wild vertebrates. *Biol. Lett.* **2012**, *8*, 983–985. [CrossRef]
22. Booth, W.; Johnson, D.H.; Moore, S.; Schal, C.; Vargo, E.L. Evidence for viable, non-clonal but fatherless *Boa constrictors*. *Biol. Lett.* **2011**, *7*, 253–256. [CrossRef]
23. Booth, W.; Schuett, G.W. Molecular genetic evidence for alternative reproductive strategies in North American pitvipers (Serpentes: Viperidae): Long-term sperm storage and facultative parthenogenesis. *Biol. J. Linn. Soc.* **2011**, *104*, 934–942. [CrossRef]
24. Dubach, J.; Sajewicz, A.; Pawley, R. Parthenogenesis in the *Arafuran filesnake* (*Acrochordus arafurae*). *Herpetol. Nat. Hist.* **1997**, *5*, 11–18.
25. Schuett, G.; Fernandez, P.J.; Gergits, W.F.; Casna, N.J.; Chiszar, D.; Smith, H.M.; Mitton, J.B.; Mackessy, S.P.; Odum, R.A.; Demlong, M.J. Production of offspring in the absence of males: Evidence for facultative parthenogenesis in bisexual snakes. *Herpetol. Nat. Hist.* **1997**, *5*, 1–10.
26. Groot, T.V.; Bruins, E.; Breeuwer, J.A. Molecular genetic evidence for parthenogenesis in the burmesee python, *Python molurus bivittatus*. *Heredity* **2003**, *90*, 130–135. [CrossRef]
27. Auliya, M.; Hofmann, S.; Segniagbeto, G.H.; Assou, D.; Ronfot, D.; Astrin, J.J.; Forat, S.; Koffivi, K.; Ketoh, G.; D'Cruze, N. The first genetic assessment of wild and farmed ball pythons (Reptilia, Serpentes, Pythonidae) in southern Togo. *Nat. Conserv.* **2020**, *38*, 37–59. [CrossRef]
28. Hunter, M.E.; Hart, K.M. Rapid microsatellite marker development using next generation pyrosequencing to inform invasive Burmese python—*Python molurus bivittatus*—Management. *Int. J. Mol. Sci.* **2013**, *14*, 4793–4804. [CrossRef]
29. Nussbaum, R.A. The brahminy blind snake (*Ramphotyphlops braminus*) in the Seychelles Archipelago: Distribution, variation, and further evidence for parthenogenesis. *Herpetologica* **1980**, *16*, e0252049.

30. Cacioppo, J.A.; Perry, S.M.; Rockwell, K.; Mitchell, M.A. A Survey of Husbandry and Breeding Techniques in the Ball Python (*Python regius*) Pet Trade. *J. Herpetol. Med. Surg.* **2020**, *31*, 25–35. [CrossRef]
31. Bertocchi, M.; Pelizzone, I.; Parmigiani, E.; Ponzio, P.; Macchi, E.; Righi, F.; Di Girolamo, N.; Bigliardi, E.; Denti, L.; Bresciani, C.; et al. Monitoring the reproductive activity in captive bred female ball pythons (*P. regius*) by ultrasound evaluation and noninvasive analysis of faecal reproductive hormone (progesterone and 17β-estradiol) metabolites trends. *PLoS ONE* **2018**, *13*, e0199377. [CrossRef]
32. Harrington, L.A.; Green, J.; Muinde, P.; Macdonald, D.W.; Auliya, M.; D'Cruze, N. Snakes and ladders: A review of ball python production in West Africa for the global pet market. *Nat. Conserv.* **2020**, *41*, 1–24. [CrossRef]
33. Siegel, D.S.; Miralles, A.; Chabarria, R.E.; Aldridge, R.D. Female reproductive anatomy: Cloaca, oviduct, and sperm storage. *Reprod. Biol. Phylogeny Snakes* **2011**, *9*, 347–409. [CrossRef]
34. Sever, D.M.; Hamlett, W.C. Female sperm storage in reptiles. *J. Exp. Zool.* **2002**, *292*, 187–199. [CrossRef] [PubMed]
35. Levine, B.A.; Schuett, G.W.; Booth, W. Exceptional long-term sperm storage by a female vertebrate. *PLoS ONE* **2021**, *16*, e0252049. [CrossRef]
36. Lampert, K.P. Facultative parthenogenesis in vertebrates: Reproductive error or chance? *Sex Dev.* **2008**, *2*, 290–301. [CrossRef] [PubMed]
37. Gamble, T.; Castoe, T.A.; Nielsen, S.V.; Banks, J.L.; Card, D.C.; Schield, D.R.; Schuett, G.W.; Booth, W. The discovery of XY sex chromosomes in a boa and python. *Curr. Biol.* **2017**, *27*, 2148–2153. [CrossRef]

animals

MDPI

Article

Chromosome Instability in Pony of Esperia Breed Naturally Infected by Intestinal Strongylidae

Emanuele D'Anza [1,†], Francesco Buono [1,†], Sara Albarella [1,*], Elisa Castaldo [1], Mariagiulia Pugliano [1], Alessandra Iannuzzi [2], Ilaria Cascone [1], Edoardo Battista [3], Vincenzo Peretti [1] and Francesca Ciotola [1]

[1] Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Via Delpino 1, 80137 Naples, Italy
[2] National Research Council (CNR), Institute of Animal Production System in Mediterranean Environment (ISPAAM), Piazzale E. Fermi, 1, 80055 Portici, Italy
[3] Independent Researcher, Via Rampa 6, 03038 Roccasecca, Italy
* Correspondence: sara.albarella@unina.it; Tel.: +39-0812536502
† These authors contributed equally to this work.

Simple Summary: Intestinal parasites are among the main causes of hidden economic losses in livestock farming. This study reports the results of chromosome instability analyses in Esperia ponies with different intestinal strongyles fecal egg counts. Interestingly, animals with higher fecal egg counts showed increased levels of chromosome instability. If this condition is confirmed in other horse breeds and livestock species, it will be important to understand the causes in order to implement therapeutic strategies for the management of intestinal parasites.

Abstract: The Pony of Esperia is an Italian autochthonous horse breed reared in the wild on the Aurunci and Ausoni Mountains. Currently, it is considered an endangered breed, as its population consists of 1623 animals. It is therefore essential to identify all aspects that can improve the management and economy of its breeding, favoring its diffusion. In this paper, the effects of intestinal strongyle infection on the chromosome stability of peripheral blood lymphocytes (PBLs) was evaluated through aneuploidy and chromosome aberration (gap, chromatid and chromosome breaks, and the number of abnormal cells) test. Statistical difference in the mean values of aneuploidy, cells with chromosome abnormalities, and chromosome and chromatid breaks were observed between ponies with high fecal egg counts (eggs per gram > 930) and those with undetectable intestinal strongylosis. The causes of this phenomenon and possible repercussions on the management of Pony of Esperia are discussed in the paper.

Keywords: Pony of Esperia; chromosome instability (CIN); intestinal strongylosis; eggs per gram (EPG); chromosome aberrations (CAs)

Citation: D'Anza, E.; Buono, F.; Albarella, S.; Castaldo, E.; Pugliano, M.; Iannuzzi, A.; Cascone, I.; Battista, E.; Peretti, V.; Ciotola, F. Chromosome Instability in Pony of Esperia Breed Naturally Infected by Intestinal Strongylidae. *Animals* **2022**, *12*, 2817. https://doi.org/10.3390/ani12202817

Academic Editor: Markku Saastamoinen

Received: 25 August 2022
Accepted: 14 October 2022
Published: 18 October 2022

1. Introduction

The Pony of Esperia (Figure 1) is an autochthonous horse breed reared in the province of Frosinone (Lazio, Central Italy). It is characterized by a morello coat, sometimes with socks and head star, and a very thick mane and tail. The head is short and conical with a straight profile; the neck is proportionate and not excessively muscular; the robust shoulder is well-attached to the trunk; the withers is pronounced; the back is inclined; the chest is developed and muscular; the thorax is shallow; and the limbs are robust. It has a maximum withers height of 138 cm for males and 132 cm for females with a maximum live weight of 350 Kg. It originated in the area of the Aurunci and Ausoni Mountains and underwent crossbreeding along successive generations with Arabian horses of Nedjad origin. The selective pressure of the mountainous environment in which it developed determined its small size, the ability to live in hostile environments, and its rusticity. This horse was used for different purposes such as the transport of materials or people, riding,

and mule production. The Pony of Esperia is currently bred as a saddle animal and for riding competitions.

Figure 1. (**a**) Pony of Esperia. (**b**) Grazing herd of Pony of Esperia.

This breed is reared in the wild and it stays in the pastures throughout the year. No food supplementations are administered, and pharmacological treatments are performed only when strictly necessary.

Currently, there are 1623 individuals enrolled in the studbook, so it is included in the breeds with limited diffusion to be safeguarded. For this purpose, it is necessary to improve our knowledge about their geno-morphofunctional characteristics and to highlight the conditions that make its farming more difficult and economically burdensome. The main breeding issues to which this breed is exposed are attacks by predators (which usually affect younger individuals or those in a precarious state of health), infectious diseases transmitted by other wild animals or transmitted from endo-ectoparasites. A study performed in 2006 on 230 individuals belonging to 33 families, showed the presence of polyparasitism in this breed. Intestinal Strongyles, *Parascaris* spp., *Oxyuris equi*, *Anoplocephala* spp., ticks, flies, and *Gasterophilus* spp. larvae have been found in all groups, suggesting that despite this breed being considered rustic and disease-resistant it needs parasitosis control plans [1].

Intestinal parasites are often responsible for delays in growth and the worsening of athletic performance, causing hidden economic losses in breeding [2]. Since they are often clinically asymptomatic, in livestock preventive therapies are carried out at specific times of the year, but this is not the case in Pony of Esperia breeding.

The aim of this study was to evaluate the effect of different degrees of natural intestinal strongyles infection on peripheral blood lymphocytes (PBLs) chromosome instability (CIN) in Pony of Esperia using aneuploidy and chromosome aberration (CA) tests [3]. Both tests have been used in humans and animals to evaluate the mutagenetic effects of environmental pollutants [4–6], drug and food supplements [7,8]; correlations among congenital malformations and chromosome stability [9–11], chromosome stability differences within different breeds of the same species [12], and the effects of micro- and macronutrient deficiencies [13].

PBLs are the ideal cells for CIN evaluation in an individual since they circulate throughout the body, thus being exposed to all possible risk conditions and thus representing an early and easy marker to analyze after in vitro cultivation [14].

CIN is a type of genomic instability with an increase in the numerical and structural alterations in chromosomes, and it is due to an increase in DNA damages, the malfunction of DNA damage repair mechanisms, or both [15]. The increase in CIN, as well as being a risk factor for the development of cancers, represents an index of altered homeostasis that negatively influences the well-being of the individual. Moreover, in livestock an increase in CIN has been negatively correlated to fertility and reproduction [16].

To the best authors knowledge, this is the first time that a relation between aneuploidy and CAs and the degree of intestinal parasitic infection was investigated in horse species.

2. Materials and Methods

2.1. Animals

For this study, fifty female Ponies of Esperia, aged between 3 and 20 years, were enrolled. All animals belonged to the same farm located in the province of Frosinone (Lazio Region) and were reared under the same conditions. All individuals were sampled twice (D0, before the treatment, and D14, 14 days after the treatment) for blood and feces to perform karyotype, aneuploidy, and CA tests in PBLs and fecal egg counts (FEC), respectively. At both sampling times, a clinical evaluation was performed, and the body condition scores (BCS) were determined using a five-point system (1 = poor, 2 = moderate, 3 = ideal, 4 = fat, and 5 = obese) [17] by the same investigator, and all animals were healthy, with a BCS = 3.

2.2. Coprological Analysis

Individual fecal samples were collected from all ponies involved in the study, and according to general recommendations proposed by Nielsen et al. [18], feces were taken directly from the rectum of each animal. Individual fecal egg counts (FECs) were performed for all ponies before the start of the trial (D0) and at 14 days post-treatment (D14) using a special modification of the McMaster method with a lower detection limit of 10 eggs per gram (EPG) using a Sheather's saturated sugar solution with a specific gravity of 1.250 as a flotation medium [19]. Based on the morphological identification [20], each egg was classified as belonging to intestinal strongyles, *Parascaris* spp., *Strongyloides westeri*, *Oxyuris equi*, or *Anoplocephala* spp.

2.3. Anthelmintic Treatment

After the first collection, the ponies were divided in two groups of 25 animals that were homogeneous per age and fecal egg count: in the treatment group (T group), animals were treated with fenbendazole (FBZ) at the horse dose rate (7.5 mg/kg BW, Panacur Oral Paste, MSD Animal Health, Walton, Milton Keynes, UK), and control group (C group) animals were left untreated. Two weeks after the treatment (day 14) the blood and feces of all animals were resampled to verify changes in aneuploidy, CA, and parasitic infection. To determine the anthelmintic efficacy of FBZ, the arithmetic mean (AM) of EPG was calculated 14 days post-treatment, and the percent efficacy (%) was considered in terms of a fecal egg count reduction (FECR) test using the formula: FECR = [(AM FEC PRE-TREATMENT − AM FEC POST-TREATMENT)/AM FEC PRE-TREATMENT] × 100, according to the American Association of Equine Practitioners (AAEP) guidelines [21]. The cut-off values used to interpret the results of the FECRT were the following: efficacy > 95%, suspected resistance 90–95%, and resistance < 90% [21]. Microsoft Office Excel 2010 software was used for data recording, and FEC reductions, expressed as percentages with 95% confidence intervals, were calculated using the RESO FECRT analysis program, version 4 [22], for Excel. The simultaneous finding of a lower confidence limit (LCL) below 90% [23] and a mean percentage of FECR below 90% [21] was indicative of resistance; if only one of these two criteria was present, resistance was suspected.

2.4. Cytogenetic Analyses

Cell cultures for chromosome isolation were set up as reported by Ciotola et al. [24]. Briefly, peripheral blood (1 mL) was cultured at 37.5 °C in RPMI medium and enriched with fetal calf serum (10%), L-glutamine (1%), and lectin (1.5%) for 48 h. Cells were harvested after colcemid (0.3 lg/mL) treatment for 1 h and given a hypotonic treatment (KCl 0.5%) and three fixations in methanol–acetic acid (3:1), the third occurring overnight. Three drops of cell suspension were air-dried on cleaned and wet slides that were stained a day later with acridine orange (0.1% in a phosphate buffer, pH 7.0) for 3 min, washed in tap and distilled water, and mounted in the same phosphate buffer. The slides were observed 24 h after staining or later (1 week). At least 10, 100, and 50 cells per animal were examined from slides of normal cultures to perform conventional karyotype, aneuploidy, and chromosome

aberration (CA) tests, i.e., gaps, chromatid and chromosome breaks, and the percentage of abnormal cells (abnormal cells are those with at least one chromatid or chromosome break) (Figure 2) [8,25]. All metaphase plates were observed under a fluorescence Nikon Eclipse 80i microscope, captured with a Nikon Sight DS-5M digital camera, transferred to a PC, and later processed with an image analysis software by two cytogeneticists.

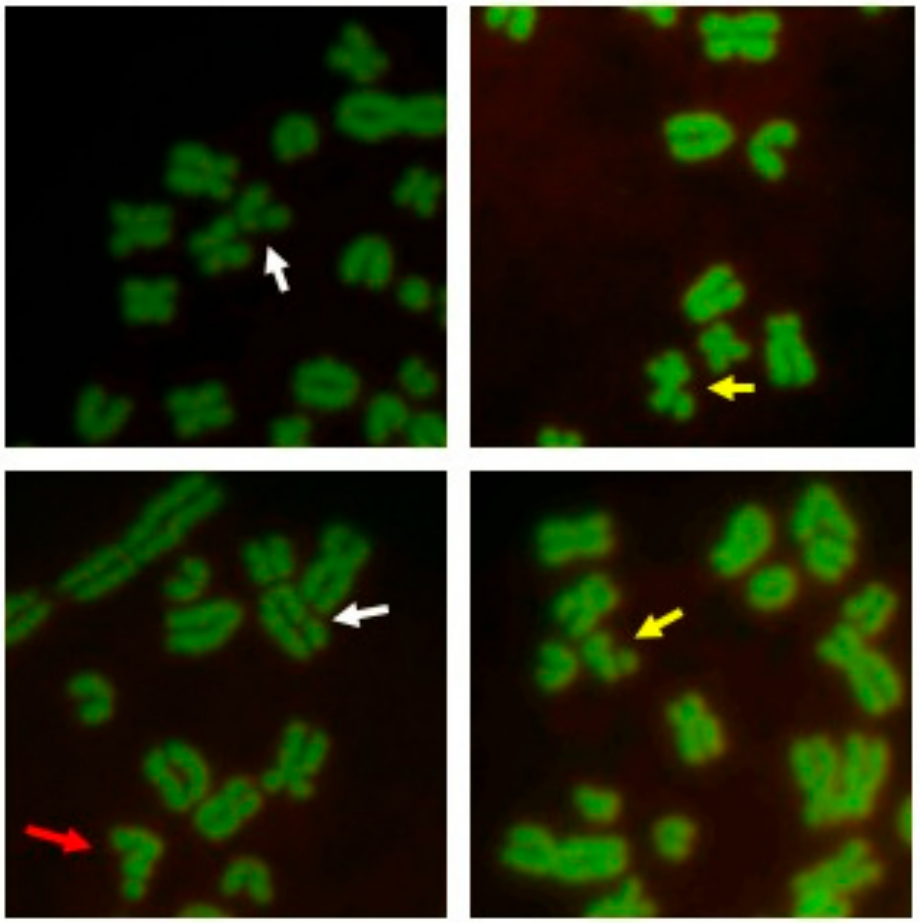

Figure 2. Details of chromosome metaphase plates of Pony of Esperia. The white arrow indicates a gap, the yellow arrow indicates a chromatid break, and the red arrow indicates a chromosome break.

2.5. Statistical Analyses

Microsoft Office Excel 2010 software was used for data recording then IBM SPSS for Windows software package version 22.0 (SPSS Inc., Chicago, IL, USA) was used for statistical analyses. The Esperia ponies were divided into five groups: group C + T included all the animals at day 0, groups C_{D0} and C_{D14} referred to the control group on day 0 and day 14, respectively, and groups T_{D0} and T_{D14} referred to the treated animals on day 0 and day 14, respectively. A statistical subanalysis was performed by dividing the groups C + T and T_{D0} according to the EPG and then according to age of the individuals (up to 6 years and equal or higher to 6 years). The parasitic burden ranges of the groups were chosen by making multiple comparisons to verify if there were ranges within which the CAs increased or decreased significantly. Age groups were established according to Wójcik and Smalec [25].

Student's t-test was used to compare the structural percentage and cells with the CA percentage in all groups. The independent-sample *t*-test (Mann–Whitney test) was used to compare the means of the quantitative variables in the groups [26]. A Spearman correlation was performed between FEC and CAs.

3. Results

3.1. Coprological Analysis and Anthelmintic Efficacy

Intestinal strongyle eggs were found in all tested animals. At the start of the study, the mean EPG count was 992.20 ± 443.75. The mean EPG values in the T and C groups were 1096.80 ± 424.31 and 887.60 ± 446.31, respectively, and no statistical differences were observed between the two group (t = 1.6986; $p = 0.0959$). For all animals, the dose was administered carefully, and no adverse reactions were observed in any of the treated ponies. FBZ was effective in reducing FECs at 14 days after treatment, showing an FECRT = 100%.

3.2. Cytogenetic Analyses

All investigated animals showed a normal karyotype, thus excluding the presence of congenital numerical and structural chromosome abnormalities that could alter CIN. Thus, the identified aneuploidy and chromosomal abnormalities observed in the metaphases of the analyzed horses did not cause birth defects.

With regards to aneuploidy and the CA assay: for seven animals at D0 and another seven animals at D14, it was not possible to analyze enough metaphases.

As it has been reported that chromosomal stability in horses is influenced by age [25], to verify if this parameter significantly influenced the values of aneuploidy and CAs in the studied population, the T_{D14} group (animals negative for intestinal parasites) was divided into two groups: $T_{D14} < 5$ ($n = 9$) and $T_{D14} > 6$ ($n = 9$), comprising, respectively, individuals aged less than or equal to 5 years and individuals older than or equal to 6 years. Statistically significant differences were observed for gaps and aneuploidy (Table 1).

When considering aneuploidy, it was decreased in all groups after the first collection in a statistically significant way, while all CA values were increased in group C_{D14} and decreased in T_{D14}.

When comparing groups divided according to EPG (Table 2), statistically significant differences were found for chromatid breaks, chromosome breaks, and CAs excluding gaps between groups $1T_{D0}$ and $2T_{D0}$, with all values being higher in $2T_{D0}$. In the comparison between $1T_{D0}$ and T_{D14}, a statistically significant difference was found only for gaps, while in the comparison between $2T_{D0}$ and T_{D14} all parameters showed a statistically significant differences. Statistically significant higher mean values for all CA parameters were also observed in group 3C + T when compared to T_{D14} and for chromatid breaks, chromosome breaks, and CAs excluding gaps in 3C + T when compared to groups $1T_{D0}$ and 1C + T.

When animals grouped according to EPG and age (up to 5 years and older than 6 years) were compared, a statistical difference was observed only between the aneuploidy percentage and gaps of animals with high EPG (930–1970).

Finally, the Spearman correlation test showed a negative correlation between EPG and gaps ($r = -0.07$; $p < 0.01$) and positive correlations between EPG and chromosome breaks ($r = 0.05$; $p < 0.01$) and EPG and CAs excluding gaps ($r = 0.33$; $p < 0.05$).

Table 1. Percentages of aneuploidy and abnormal cells and mean values and standard deviations of CA in groups C + T (all animals before the treatment), C_{D0} (untreated animals at T0), T_{D0} (treated animals before the treatment), C_{D14} (untreated animals at T14), T_{D14} (treated animals at T14), $T_{D14} < 6$ (treated animals at T14 up to 6 years), and $T_{D14} > 6$ (treated animals at T14 older than 6 years).

Group	N of Animals	Age		EPG		Aneuploidy (2n ≠ 64)	Abnormal Cells	Gaps	Chromatid Breaks	Chromosome Breaks	CAs Excluding Gaps
		Mean ± SD	Range	Mean ± SD	Range	%	%	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
C + T	43	7.63 ± 5.02	3–20	980.24 ± 471.75	260–1970	16.76 ^{a,A}	6.33	0.71 ± 0.93	0.06 ± 0.25 ^{a}	0.01 ± 0.11 ^{a}	0.07 ± 0.28 ^{A}
C_{D0}	20	8.00 ± 6.04	3–20	904.44 ± 536.34	260–1970	15.44 ^{a,b,A}	4.70	0.72 ± 0.98	0.04 ± 0.22 ^{a,A}	0.01 ± 0.10	0.05 ± 0.25 ^{a}
T_{D0}	23	7.45 ± 4.06	3–20	1039.57 ± 389.47	400–1970	17.90 ^{a,A}	7.74	0.70 ± 0.88	0.07 ± 0.27 ^{B}	0.01 ± 0.12	0.08 ± 0.30 ^{A,b}
C_{D14}	13	7.75 ± 6.25	3–20	1179.00 ± 801.82	110–2760	11.30 ^{c,B}	7.20	0.79 ± 0.91 ^{A}	0.07 ± 0.28 ^{B}	0.01 ± 0.12	0.08 ± 0.31 ^{A,b}
T_{D14}	23	7.45 ± 4.06	3–20	0	0	13.32 ^{b,c}	6.64	0.61 ± 0.74	0.04 ± 0.19 ^{b,A}	0.00 ± 0.07 ^{b}	0.04 ± 0.2 ^{B}
$T_{D14} < 6$	12	3.91 ± 0.95	3–5	0	0	11.25 ^{c,B}	6.31	0.77 ± 0.80 ^{A}	0.04 ± 0.18 ^{b,A}	0.00 ± 0.00 ^{a}	0.04 ± 0.18 ^{B}
$T_{D14} > 6$	11	10.25 ± 4.91	6–20	0	0	16.05 ^{b,A}	7.45	0.49 ± 0.66 ^{B}	0.04 ± 0.21 ^{b,A}	0.00 ± 0.06 ^{a}	0.05 ± 0.22 ^{a}

^{a,b,c} = $p < 0.05$; ^{A,B} = $p < 0.005$.

Table 2. Percentages of aneuploidy and abnormal cells and mean values and standard deviations of chromosome abnormalities in ponies grouped according to EPG and age at D0 and compared with the T_{D14} group.

Group	N of Animals	Age		EPG Range	Aneuploidy (2n ≠ 64)	Abnormal Cells	Gaps	Chromatid Breaks	Chromosome Breaks	CAs Excluding Gaps
		Mean	Range		%	%	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
$1T_{D0}$	11	7.08 ± 4.32	3–16	400–800	17.39	5.69	0.87 ± 0.88 ^{A}	0.05 ± 0.23 ^{A}	0.00 ± 0.07 ^{A}	0.05 ± 0.31 ^{A,a}
$2T_{D0}$	9	8.23 ± 5.43	3–20	950–1970	14.47	10.22	0.81 ± 0.86 ^{A}	0.08 ± 0.29 ^{B}	0.03 ± 0.16 ^{B}	0.11 ± 0.33 ^{B}
1C + T	12	8.4 ± 6.12	3–20	260–640	18.62	5.20	0.74 ± 0.96	0.05 ± 0.23 ^{A}	0.00 ± 0.06 ^{A}	0.05 ± 0.24 ^{A,a}
2C + T	13	6.17 ± 3.24	3–12	690–930	12.26 ^{a}	6.92	0.76 ± 0.94	0.07 ± 0.27	0.01 ± 0.10	0.08 ± 0.30 ^{b}
3C + T	18	9.00 ± 6.14	3–20	950–1970	17.85	10.89	0.81 ± 0.88 ^{A}	0.08 ± 0.30 ^{B}	0.04 ± 0.19 ^{B}	0.12 ± 0.35 ^{B}
1–2C + T < 6	12	3.82 ± 0.94	3–5	260–930	19.13 ^{b}	10.00	0.76 ± 0.94	0.06 ± 0.27	0.01 ± 0.09	0.07 ± 0.29
1–2C + T > 6	11	10.24 ± 4.15	6–20	260–930	18.41	9.64	0.80 ± 0.95	0.06 ± 0.25	0.01 ± 0.08	0.07 ± 0.27
3C + T < 6	9	3.40 ± 0.80	3–5	950–1970	12.30 ^{a}	10.64	1.12 ± 0.98 ^{A}	0.06 ± 0.26	0.05 ± 0.21 ^{B}	0.11 ± 0.34 ^{B}
3C + T > 6	9	11.00 ± 4.79	6–20	950–1970	16.52	11.67	0.72 ± 0.83 ^{B}	0.08 ± 0.31 ^{B}	0.03 ± 0.17 ^{B}	0.11 ± 0.34 ^{B}
T_{D14}	23	7.45 ± 4.06	3–20	0	13.32	6.64	0.61 ± 0.74 ^{B}	0.04 ± 0.19 ^{A}	0.00 ± 0.07 ^{A}	0.04 ± 0.21 ^{A,a}

^{a,b} = $p < 0.05$; ^{A,B} = $p < 0.005$.

4. Discussion

The Pony of Esperia is an endangered autochthonous breed reared in the Aurunci and Ausoni Italian Mountains. The implementation of safeguard plans for this breed is essential not only to avoid the loss of biodiversity but also to avoid the abandonment of marginal areas. For this purpose, it appears important to highlight the causes that can worsen the profitability of this breed and to verify the treatments that can improve its productivity.

One of the main causes of hidden economic losses in animal farming is parasitic infections of digestive tract. In fact, they can induce states of subclinical malnutrition due to the alteration of the intestinal absorption capacity as well as the actual subtraction of nutrients by parasites.

In this pony population, anthelmintic treatments are carried out without performing a coprological diagnosis once per year, and this anthelmintic treatment scheme is quite different from those reported in horses in Italy [27] but is very similar to that reported in the Italian donkey population [28]. Contrary to the diffusion of anthelmintic resistance worldwide [29], anthelmintic treatment with fenbendazole has been shown to be highly effective in this wild pony population, and for this reason, as reported for donkeys [30], ponies of Esperia can be considered an animal population in refugia.

With regards to the effect of intestinal strongyles on chromosome stability in PBLs, the aneuploidy and CA tests have provided interesting results.

After dividing group T_{D14} (animals with EPG = 0) into two groups based on age (up to 5 years and equal to or older than 6 years), a statistically significant difference was observed between aneuploidy and gap. It is interesting to note that, while the gaps were higher in younger ponies, a reduction was observed in the older ones in favor of the percentage of aneuploid cells that instead was higher in older animals. This result is congruent with what was observed by Wojcik and Smalec [25], who found that in horses the mean value of SCEs was significantly different between horses under 6 years of age and horses older than 6 years, indicating that the age of the investigated animals also represents a reference parameter for sampling when studying CIN by aneuploidy and CAs but only when considering gaps. In fact, abnormal cells percentages, chromatid and chromosome breaks, and CAs excluding gap are not affected by age.

The main finding is that the mean values of chromosome breaks and CAs excluding gaps significantly increase in ponies with high fecal egg counts. Another interesting result is that a statistically significant higher value of CAs was found in animals with EPG > 930 when compared to animals with EPG < 930. In this regard, it has been seen that, in equids, the number of parasites present could be decisive in evaluating the state of health of an individual, for which EPG < 500 represents a low infection, 500 < EPG < 1000 represents a moderate infection, and EPG > 1000 represents a high infection [31,32]. The only parameter in which an inconsistency was observed is aneuploidy, which was higher in subjects with 260 < EPG < 640 than in animals with 690 < EPG < 930. However, the differences were not statistically significant.

When animals were grouped by EPG and then by age, statistically significant differences were only observed in the mean number of gaps in individuals with high EPG (950–1970).

The data reported here show that intestinal parasites affect chromosomal stability as evaluated by aneuploidy and CAs. The studies published to date on the application of chromosomal stability tests in different animal species show correlations with age and breed, while correlations with any parasitic condition have never been considered. This study is therefore the first to report this effect, but the data are limited to aneuploidy and CA tests. It therefore appears necessary to verify whether other tests, such as SCE, micronuclei, the comet assay, and fragile sites, can also be affected by intestinal parasites or other types of endoparasitosis.

The CA test detects exposure to substances that cause DNA strand breaks or a deficiency of substances that are involved in DNA repair mechanisms but does not allow the identification of the clastogen class or micronutrient deficiency [33]. Thus, it is possible to

hypothesize that increased CAs in animals with high EPG could be due to a subclinical deficiency of micro- and macronutrients linked to the action of parasites present in the intestine, which in addition to subtracting nutrients, cause alterations in the composition of the microbiota [34] (fundamental in these species to produce micro- and macronutrients) and to the structure of the intestinal barrier, with consequences for the absorption ability. Another possibility is that the alterations in the intestinal barrier of the microbiota and the inflammatory processes induced by parasites allow the passage of toxic substances into the circulation that would normally be eliminated with the feces and that have a negative effect on PBLs. Both situations may cause decreases in sport performance and fertility, which in both cases, become possible causes of reduced earnings with significant impacts on the profitability of the breeding of native breeds such as the Pony of Esperia.

These hypotheses, if confirmed through specific dosage analyses of micro- and macronutrients or of other substances at the serum level, would suggest the intervention, in subjects with EPG > 930, with appropriate food supplements. Moreover, it would be interesting to evaluate in other horse breeds and in other livestock species the impact on CIN of intestinal strongyle FEC, verifying the minimum EPG count at which it is appropriate to carry out pharmacological treatments, possibly associated with food supplement administration. These data would be useful for breeds raised in the wild, for the semi-extensive technique in which carrying out pharmacological treatments is not particularly easy, and for animals in intensive farming in which the improper use of pharmacological molecules causes a considerable environmental impact. Furthermore, in native breeds these data acquire further value since the low profitability of breeding involves an accurate evaluation of the cost/benefit ratios of any treatment that is selected.

5. Conclusions

According to the results observed in this study, it is possible to speculate that in the wildly reared Pony of Esperia breed: (1) an anthelmintic treatment with fenbendazole is highly effective, (2) a high FEC (>930 EPG) causes an increase in CAs that is reduced after only 14 days of drug treatment, and (3) in a cost–benefit assessment, the treatment of parasitosis could be useful in subjects with FEC > 930. Finally, an important aspect emerges that should be verified with future studies: individuals with FEC > 930 could benefit from food supplementation with micro- and macronutrients.

Author Contributions: Conceptualization, F.C. and A.I.; Methodology, S.A., F.B. and A.I.; Formal Analysis, E.D., E.C., I.C., M.P., E.B. and F.B.; Resources, F.C. and E.B.; Funding Acquisition, V.P. and F.C.; Data Curation, E.D., S.A., F.B. and I.C.; Writing—Original Draft Preparation, E.D., I.C., M.P., E.C. and E.B.; Writing—Review and Editing, S.A., F.B. and A.I.; Supervision, F.C. and V.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was carried out following the recommendations of the European Council Directive (86/609/EEC) on the protection of animals. Ethical approval was not required in this study, as fecal and blood sampling were performed by the vet as part of a routine clinical visit.

Informed Consent Statement: Informed consent was obtained from the owner of the animals.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors wish to thank Flavio Luongo for his technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rinaldi, L.; Musella, V.; Santaniello, A.; Santaniello, M.; Carbone, S.; Battista, E.; Cringoli, G. Mappe parassitologiche del pony di Esperia. In *Il Pony di Esperia*, 1st ed.; Ciociariaturismo: Frosinone, Italy, 2009; pp. 126–156.
2. Love, S.; Murphy, D.; Mellor, D. Pathogenicity of cyathostome infection. *Vet. Parasitol.* **1999**, *85*, 113–121. [CrossRef]
3. Peretti, V.; Ciotola, F.; Albarella, S.; Russo, V.; Di Meo, G.; Iannuzzi, L.; Roperto, F.; Barbieri, V. Chromosome fragility in cattle with chronic enzootic haematuria. *Mutagenesis* **2007**, *22*, 317–320. [CrossRef] [PubMed]
4. Di Meo, G.P.; Perucatti, A.; Genualdo, V.; Caputi-Jambrenghi, A.; Rasero, R.; Nebbia, C.; Iannuzzi, L. Chromosome fragility in dairy cows exposed to dioxins and dioxin-like PCBs. *Mutagenesis* **2011**, *26*, 269–272. [CrossRef]
5. Genualdo, V.; Perucatti, A.; Iannuzzi, A.; Di Meo, G.P.; Spagnuolo, S.M.; Caputi-Jambrenghi, A.; Coletta, A.; Vonghia, G.; Iannuzzi, L. Chromosome fragility in river buffalo cows exposed to dioxins. *J. Appl Genet.* **2012**, *53*, 221–226. [CrossRef] [PubMed]
6. Perucatti, A.; Di Meo, G.P.; Albarella, S.; Ciotola, F.; Incarnato, D.; Caputi Jambreghi, A.; Peretti, V.; Vonghia, G.; Iannuzzi, L. Increased frequencies of both chromosome abnormalites and SCEs in two sheep flocks exposed to high dioxin levels during pasturages. *Mutagenesis* **2006**, *21*, 67–75. [CrossRef]
7. Ames, B.N. DNA damage from micronutrient deficiencies is likely to be a major cause of cancer. *Mutat. Res.* **2001**, *475*, 7–20. [CrossRef]
8. Savage, J.R. Classification and relationships of induced chromosomal structural changes. *J. Med. Genet.* **1975**, *13*, 103–122. [CrossRef]
9. Albarella, S.; Ciotola, F.; Dario, C.; Iannuzzi, L.; Barbieri, V.; Peretti, V. Chromosome instability in Mediterranean Italian buffaloes affected by limb malformation (transversal hemimelia). *Mutagenesis* **2009**, *24*, 471–474. [CrossRef]
10. Macrì, F.; Ciotola, F.; Rapisarda, G.; Lanteri, G.; Albarella, S.; Aiudi, G.; Liotta, L.; Marino, F. A rare case of simple syndactyly in a puppy. *J. Small Anim. Pract.* **2014**, *55*, 170–173. [CrossRef]
11. Peretti, V.; Ciotola, F.; Albarella, S.; Paciello, O.; Dario, C.; Barbieri, V.; Iannuzzi, L. XX/XY chimerism in cattle: Clinical and cytogenetic studies. *Sex. Dev.* **2008**, *2*, 24–30. [CrossRef]
12. Ciotola, F.; Albarella, S.; Scopino, G.; Carpino, S.; Monaco, F.; Peretti, V. Crossbreeding effect on genome stability in pig (*Sus scrofa scrofa*). *Folia Biol.* **2014**, *62*, 23–28. [CrossRef] [PubMed]
13. Ames, B.N. Micronutrient deficiencies: A major cause of DNA damage. *Ann. N. Y. Acad Sci.* **1999**, *889*, 87–106. [CrossRef] [PubMed]
14. Johannes, C.; Obe, G. Chromosomal aberration test in human lymphocytes. *Methods Mol. Biol.* **2013**, *1044*, 165–178. [CrossRef]
15. Siri, S.O.; Martino, J.; Gottifredi, V. Structural Chromosome Instability: Types, Origins, Consequences, and Therapeutic Opportunities. *Cancers* **2021**, *13*, 3056. [CrossRef] [PubMed]
16. Wójcik, E.; Sokól, A. Assessment of chromosome stability in boars. *PLoS ONE* **2020**, *15*, e0231928. [CrossRef]
17. Carroll, C.L.; Huntington, P.J. Body condition scoring and weight estimation of horses. *Equine Vet. J.* **1988**, *20*, 41–45. [CrossRef]
18. Nielsen, M.K.; Vidyashankar, A.; Andresen, U.; De Lisi, K.; Pilegaard, K.; Kaplan, R. Effects of fecal collection and storage factors on strongylid egg counts in horses. *Vet. Parasitol.* **2010**, *167*, 55–61. [CrossRef]
19. Lester, H.; Matthews, J.B. Faecal worm egg count analysis for targeting anthelmintic treatment in horse: Points to consider. *Equine Vet. J.* **2014**, *46*, 139–145. [CrossRef]
20. Zajac, A.M.; Conboy, G.A. *Veterinary Clinical Parasitology*, 8th ed.; Wiley Blackwell Press: Oxford, UK, 2011; 354p.
21. Nielsen, M.K.; Mittel, L.; Grice, A.; Erskine, M.; Graves, E.; Vaala, W.; Tully, R.C.; French, D.D.; Bowman, R.; Kaplan, R.M. American Association of Eequine Practioners Parasite Control Guidelines. 2019. Available online: https://aaep.org/sites/default/files/Guidelines/AAEPParasiteControlGuidelines_0.pdf (accessed on 25 May 2022).
22. Wursthorn, L.; Martin, P. *Reso 2.0: Calculation for Fecal Egg Count Reduction Test (FECRT)*; Animal Health Research Laboratory—CSIRO: Canberra, Australia, 1990.
23. Tzelos, T.; Barbeito, J.; Nielsen, M.K.; Morgan, E.; Hodgkinson, J.; Matthews, J.B. Strongyle egg reappearance period after moxidectin treatment and its relationship with management factors in UK equine populations. *Vet. Parasitol.* **2017**, *237*, 70–76. [CrossRef]
24. Ciotola, F.; Albarella, S.; Pasolini, M.P.; Auletta, L.; Esposito, L.; Iannuzzi, L.; Peretti, V. Molecular and cytogenetic studies in a case of XX SRY-negative sex reversal in an Arabian horse. *Sex. Dev.* **2012**, *6*, 104–107. [CrossRef]
25. Wójcik, E.; Smalec, E. The effect of environmental factors on sister chromatid exchange incidence in domestic horse (*Equus caballus*) chromosomes. *Folia Biol.* **2013**, *61*, 199–204. [CrossRef] [PubMed]
26. Carrano, A.V.; Natarajan, A.T. International Commission for Protection Against Environmental Mutagens and Carcinogens. ICPEMC publication no. 14. Considerations for population monitoring using cytogenetic techniques. *Mutat. Res.* **1988**, *204*, 379–406. [CrossRef]
27. Veneziano, V.; Veronesi, F.; Buono, F.; Lia, P.R.; Manfredi, M.T.; Traversa, D.; Frangipane di Regalbono, A.; Papini, R.; Brianti, E.; Otranto, D.; et al. A national questionnaire survey of helminth control practices on horse in Italy—Preliminary results. In Proceedings of the 25th International Conference of the World Association for the Advancement of Veterinary Parasitology, Liverpool, UK, 16–20 August 2015; p. 366.
28. Buono, F.; Veronesi, F.; Pacifico, L.; Roncoroni, C.; Napoli, E.; Zanzani, S.A.; Mariani, U.; Neola, B.; Sgroi, G.; Piantedosi, D.; et al. Helminth infections in Italian donkeys: *Strongylus vulgaris* more common than *Dictyocaulus arnfieldi*. *J. Helminthol.* **2021**, *95*, e4. [CrossRef] [PubMed]

29. Peregrine, A.S.; Molento, M.B.; Kaplan, R.M.; Nielsen, M.K. Anthelmintic resistance in important parasites of horses: Does it really matter? *Vet. Parasitol.* **2014**, *201*, 1–8. [CrossRef]
30. Buono, F.; Roncoroni, C.; Pacifico, L.; Piantedosi, D.; Neola, B.; Barile, V.L.; Fagiolo, A.; Várady, M.; Veneziano, V. Cyathostominae Egg Reappearance Period After Treatment with Major Horse Anthelmintics in Donkeys. *J. Equine Vet. Sci.* **2018**, *65*, 6–11. [CrossRef]
31. Scala, A.; Tamponi, C.; Sanna, G.; Predieri, G.; Dessì, G.; Sedda, G.; Buono, F.; Cappai, M.G.; Veneziano, V.; Varcasia, A. Gastrointestinal Strongyles Egg Excretion in Relation to Age, Gender, and Management of Horses in Italy. *Animals* **2020**, *10*, 2283. [CrossRef]
32. Jota Baptista, C.; Sós, E.; Szabados, T.; Kerekes, V.; Madeira de Carvalho, L. Intestinal parasites in Przewalski's horses (*Equus ferus przewalskii*): A field survey at the Hortobágy National Park, Hungary. *J. Helminthol.* **2021**, *95*, e39. [CrossRef]
33. Mateuca, R.A.; Decordier, I.; Kirsch-Volders, M. Cytogenetic methods in human biomonitoring: Principles and uses. *Methods Mol. Biol.* **2012**, *817*, 305–334. [CrossRef]
34. Clark, A.; Sallé, G.; Ballan, V.; Reigner, F.; Meynadier, A.; Cortet, J.; Koch, C.; Riou, M.; Blanchard, A.; Mach, N. Strongyle Infection and Gut Microbiota: Profiling of Resistant and Susceptible Horses Over a Grazing Season. *Front. Physiol.* **2018**, *9*, 272. [CrossRef]