

Review

CSN1S1 and CSN1S2: Two Remarkable Examples of Genetically Modulated Alternative Splicing via Identification of Allele-Specific Splicing Events

Gianfranco Cosenza ¹ , Andrea Fulgione ¹ , Emanuele D'Anza ² , Sara Albarella ^{2,*} , Francesca Ciotola ² 
and Alfredo Pauciuolo ³ 

¹ Department of Agricultural Science, University of Naples Federico II, Piazza Carlo Di Borbone 1, 80055 Portici, Italy; gianfranco.cosenza@unina.it (G.C.); andrea.fulgione@unina.it (A.F.)

² Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Via Delpino 1, 80137 Naples, Italy; emanuele.danza@unina.it (E.D.); francesca.ciotola@unina.it (F.C.)

³ Department of Agricultural, Forest and Food Sciences, University of Turin, 10095 Grugliasco, Italy; alfredo.pauciuolo@unito.it

* Correspondence: sara.albarella@unina.it

Abstract

Splicing regulatory sequences are cornerstones for exon recognition. Mutations that modify them can severely compromise mRNA maturation and protein production. A wide range of mutations, including SNPs and InDels, can influence splicing regulatory signals either directly (e.g., altering canonical donor and acceptor dinucleotides) or indirectly (e.g., creating cryptic splice sites). *CSN1S1* and *CSN1S2* genes encode for the two main milk proteins, α s1 and α s2 caseins, respectively. They represent a remarkable and unique example of the possibilities for alternative splicing of individual genes, both due to the high number of alternative splices identified to date and for recognized allele-specific splicing events. To date, at least 13 alleles of *CSN1S1* originating from mutations that affect canonical splice sites have been described in *Bos taurus* (*CSN1S1* A, A1, and H), *Ovis aries* (E, H, and I), *Capra hircus* (D and G), *Bubalus bubalis* (E, F) and *Camelidae* (A, C, and D). Similarly, allele-specific splicing events have been described at the *CSN1S2* locus in *B. taurus* (*CSN1S2* D), *C. hircus* (*CSN1S2* D), *B. bubalis* (*CSN1S2* B, B1, and B2), *Equus asinus* (*CSN1S2* I B), and *Camelidae*. This review highlights that mutations affecting canonical splice sites, particularly donor sites, are significant sources of genetic variation impacting the casein production of the main dairy livestock species. Currently, a key limitation on this topic is the lack of detailed functional and proteomic studies. Future research should leverage advanced omics technologies like long-read transcriptomics and allele-resolved RNA sequencing to characterize these splicing mechanisms, guiding precision breeding strategies.

Keywords: mRNA; dairy livestock species; caseins; mutations; splice sites



Academic Editor: Dimitrios Chatzizplis

Received: 25 July 2025

Revised: 25 August 2025

Accepted: 25 August 2025

Published: 27 August 2025

Citation: Cosenza, G.; Fulgione, A.; D'Anza, E.; Albarella, S.; Ciotola, F.; Pauciuolo, A. *CSN1S1* and *CSN1S2*: Two Remarkable Examples of Genetically Modulated Alternative Splicing via Identification of Allele-Specific Splicing Events. *Genes* **2025**, *16*, 1011. <https://doi.org/10.3390/genes16091011>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Alternative splicing (or non-constitutive splicing) is an intriguing mechanism able to expand proteome complexity and gene expression. It plays a multifaceted role in several biological processes, including the modulation of gene expression [1] and the regulation of cell and tissue development and differentiation [2]. Thus, it has implications for human diseases [1], and it contributes to the evolution of complex organisms [3]. A major factor influencing the intricate mechanism of RNA maturation is the highly fragmented

architecture of genes, which accounts for the frequent occurrence of alternative splicing. Exon skipping is particularly frequent when the coding region is divided into multiple short exons, as the maturation of long primary transcripts is an intricate process requiring many successive steps [4]. Genes with multiple alternative splicing forms are expected to contain more regulatory sequences that may serve as potential target for mutations. Some reported mutations disrupt splicing by either generating new splice sites or interfering with regulatory elements. Therefore, the role of splicing mutations in gene expression and regulation may be more significant than previously believed [5].

Despite their potential significance, the number of reported mutations that alter RNA splicing remains relatively small and less investigated in the literature compared to other mutations, particularly those occurring in exons that affect a protein's functional or structural properties [6].

In this respect, very extensive non-constitutive splicing events (both alternative splicing and genetically modulated alternative splicing) were observed in the casein transcripts of various species, in particular at *CSN1S1* and *CSN1S2* loci, encoding for two main milk proteins, α s1 and α s2 caseins, respectively [4–9].

In this review, we examine sequence changes in the *CSN1S1* and *CSN1S2* genes that impact transcript structure. Although factors like spliceosomal recognition and RNA binding proteins play a role in both normal and mutation-derived alternative mRNA splicing, we specifically concentrate on genetic changes that disrupt constitutive splicing, often leading to abnormal phenotypes.

2. mRNA Splicing

A major process regulating gene expression in eukaryotes is precursor messenger RNA (pre-mRNA) splicing. It consists in the removal of intron sequences and the linear joints of remaining exons in the order in which they were in the originally transcribed DNA sequence to codify for a specific final protein [10].

The leading factor in this process is the spliceosome, an intricate molecular complex of proteins and RNA. It is composed of the small nuclear ribonucleic particles (sn-RNPs) U1, U2, U4/U6, and U5, along with other splicing factors (non-RNPs). It is responsible for the accurate removal of introns from pre-mRNA. This is achieved by identifying specific, conserved sequence elements at the exon–intron boundaries: the 5' and 3' splice sites, the branch point, and the polypyrimidine tract [11–13].

The donor splice site, at the 5' end of an intron, is characterized by the dinucleotide GT immediately after the cut point where the intron is removed from the preceding exon. Conversely, the acceptor splice site, at the 3' end of an intron, has the dinucleotide AG just before the cut point where the intron is removed from the following exon. It has been estimated that GT-AG splice site pairs represent more than 99% of all the canonical splice sites. However, several less common instances were discovered where non-canonical splice pairs were processed by related, but distinct, splicing machinery. Major non-canonical splice site combinations are GC-AG (0.56–0.69%) and AT-AC (0.05–0.09%); all sequences differing from these are non-canonical splice sites (0.02%) [14,15].

The branch point is contained within a degenerate intronic heptamer (YNYRAY common consensus sequence in animals) called “the branch point sequence”. In eukaryotes, the lariat formation site at the branch point holds an adenosine in the sequence, while the rest of the sequence is not conserved. In most introns, this motif resides between approximately 18–40 nucleotides upstream of the 3' splice site before a polypyrimidine tract. This region is rich in pyrimidines (C and U) of variable length, often interrupted by purines. It is a crucial cis-acting sequence element for intron removal, a process in which the pre-mRNA splicing factor U2AF65 plays a pivotal role. It directs the early steps in

splice site choice by recognizing the polypyrimidine tract consensus sequence located near the 3' splice site [16].

The type of splicing just described is defined as linear because a 5' splice site is connected to a downstream 3' splice site. However, it has been demonstrated that under certain conditions, the same splicing machinery and most of the splicing enhancers and silencers used for linear pre-mRNA splicing can non-collinearly splice the 5' terminus of an upstream pre-mRNA exon with the 3' terminus of a downstream exon.

This mechanism, named back-splicing, generates circular RNAs (circ-RNAs), a recently discovered class of single-stranded non-coding RNA (ncRNA). A prerequisite for the back-splicing mechanism is the spatial proximity of the corresponding splice sites, which is commonly achieved through the folding of the pre-mRNA into specific secondary structures. This unique splicing event results in a covalent bond between the 3' and 5' ends of the RNA molecule. The resulting lack of free termini renders circRNAs exceptionally stable and resistant to exoribonuclease degradation, setting them apart from their linear cognates. Back-splicing can occur in two main forms: splicing within a single exon to produce mono-exonic circRNAs, or a non-sequential joining of the 5' terminus of an upstream exon with the 3' terminus of a downstream exon, leading to multi-exonic circRNAs. Furthermore, similarly to protein-coding linear mRNAs, circRNAs can be subject to alternative splicing events.

Currently, it is estimated that circRNAs typically account for approximately 0.8–1% of linear messenger RNAs (mRNAs) on a global scale. Despite this prevalence, the overarching biological function of circRNAs remains poorly understood, and for years, they were largely dismissed as mere splicing artifacts. However, recent investigations have begun to uncover their diverse roles, including their ability to act as protein scaffolds, serve as microRNA (miRNA) sponges, and be translated into novel polypeptides [17–19].

In addition to constitutive splicing, at least two types of non-constitutive splicing have been observed: alternative splicing and allele-specific splicing.

2.1. Alternative Splicing

Alternative splicing is a ubiquitous regulatory mechanism and can be defined as a variation of constitutive splicing [12]. It is estimated that between 92% and 94% of all mRNAs in the human genome are alternatively spliced [20].

Alternative splicing is a cellular process in which exons from a single gene are joined in various combinations, generating multiple, related messenger RNA (mRNA) transcripts. These mRNA isoforms can be translated into a diverse set of proteins that possess distinct structures, functions, and/or subcellular localizations. Furthermore, certain splicing events can alter the mRNA reading frame, producing proteins with novel functions or leading to non-functional products. This mechanism significantly expands proteomic complexity from a limited number of genes [10].

There are several ways in which alternative splicing can be achieved: (a) skipping of one or more exons (also known as cassette exons); (b) activation of one or more cryptic splice sites and “exonization” of an intronic region following the recognition of canonical di-nucleotide splice acceptor and donor sites; (c) transition of a constitutive exon to an alternative exon; (d) alternative 5' splice sites (alternative donors) or alternative 3' splice sites (alternative acceptors); (e) exon shuffling, a process in which a new or a duplicated exon is inserted into a gene. The exon–intron architecture appears to be the cornerstone of alternative splicing. Long introns and short exons characterize genome architecture in vertebrates, whereas in lower eukaryotes, introns are short, and exons are long [10,15,21].

While core splice site motifs are essential, the vast majority of the regulatory information for splicing is thought to be encoded within exonic and intronic cis-regulatory

elements. These elements function as binding sites for sequence-specific RNA-binding protein factor, and they in turn either promote or suppress the recognition of proximal splice sites [22].

Furthermore, the alternative inclusion of exons with weak splice sites (weak exons) in the processed mRNA is observed. The efficiency of this inclusion is contingent upon several factors, including the activity of splicing activators or repressors, the RNA secondary structure, and the promoter structure. The RNA secondary structure, alongside the primary RNA sequence, is a critical determinant of a correct splicing process. A growing body of research also demonstrates that alternative splicing occurs in both protein-coding and non-coding genes [23].

A well-known example of a post-transcriptional consequence of alternative splicing is its intricate relationship with nonsense-mediated mRNA decay (NMD). Nonsense mutations, which introduce a premature termination codon (PTC) within an exon, can trigger these regulatory pathways. The presence of a PTC has been demonstrated to induce changes in the splicing pattern of the pre-mRNA, frequently leading to the skipping of the exon containing the mutation. This splicing alteration, coupled with the direct instability caused by NMD, results in low levels of the mutant mRNA. This mechanism serves to prevent the translation of functionally defective or potentially harmful truncated proteins [24].

Alternative splicing can also serve as a quantitative regulator of gene expression. Furthermore, it may lead to the generation of transcripts that fail to produce functional proteins. These outcomes are often attributed to variations in untranslated regions (UTRs), which can modulate mRNA stability and translational efficiency, produce truncated proteins, or change the mRNA localization, thus compromising the correct function of the transcript and/or protein. So, alternative splicing represents an important mechanism of increasing proteome complexity and regulating gene expression. An evolutionary advantage of alternative splicing is its ability to enable a single gene to produce a variety of functionally distinct mRNA transcripts that are similar but not identical. This mechanism substantially increases the coding capacity and proteomic diversity of eukaryotic genomes, as it allows for the generation of multiple protein isoforms from a single genetic locus [23,25,26].

Therefore, the full characterization of transcriptome and proteome isoforms derived by alternative splicing is essential for understanding cellular differentiation and the biological diversity observed across living organisms [27].

2.2. Allele-Specific Splicing

Consensus splicing regulatory sequences are pivotal determinants for exon recognition, and mutations that alter these sequences can profoundly impact mRNA maturation and subsequent protein synthesis. It is well established that a wide range of genetic mutations (SNPs, InDels, etc.), in both intronic and exonic regions, can disrupt these signals. This disruption may occur either directly—by altering canonical, highly conserved, donor and acceptor dinucleotides or the branch point sequence—or indirectly, through mechanisms like the creation of cryptic splice sites.

The most frequent and impactful mutations occur at the highly conserved first two nucleotides (+1/+2) of the 5' donor site and the last two nucleotides (−1/−2) of the 3' acceptor site. Notably, mutations at the donor splice site are more prevalent than those at the acceptor splice site, with a reported ratio of approximately 1.5:1 [28]. The two highly conserved splice sites exhibit significantly lower variability compared to exonic sequences, with reductions of 77% and 84% for the donor and acceptor sites, respectively. Furthermore, the intronic nucleotides immediately flanking these sites—specifically, the first eight and last twelve nucleotides at the 5' and 3' splice sites—are also more highly conserved than

the average intronic sequence [11]. When a canonical splice site is weak or a mutation activates a neighboring cryptic splice site, this alternative site can be preferentially used by the splicing machinery. The resulting outcome depends on the location of the cryptic site: if it resides within an intron, the splicing process can lead to the inclusion of an intronic fragment; conversely, if it is located within an exon, it can cause the partial or complete removal of that exon [4,14].

The allele specificity of transcript isoforms may be complete or partial. Such sequence variants often disrupt normal gene splicing, causing aberrant splicing of either a proportion or all of the transcripts produced (complete allele specificity), altering the mRNA and, consequently, the corresponding encoded protein sequences, which sometimes are aberrant or incorrectly functional. Alternatively, a single allele can produce multiple transcript isoforms (true alternative splicing). However, the relative abundance of these isoforms can exhibit significant variation between different alleles (partial allele specificity) [29–31].

Mutations occurring in the branch point sequence are rarer but can lead to exon skipping due to improper binding of the SF1 and U2 snRNP splicing proteins, resulting in disruption of the natural acceptor splicing site. Alternatively, they can cause intron retention, either partial or complete, if a new 3' splice site is created [11,12,14].

Mutations affecting the branch point adenosine or its surrounding consensus sequence can disrupt the spliceosome's ability to recognize and utilize the canonical branch point. Consequently, the splicing machinery may be prompted to utilize a nearby cryptic adenosine residue within the intron as a novel branch point. This mechanism contributes to alternative splicing events and subsequent dysregulated gene expression, which can have significant pathological consequences [11].

Proper recognition of the polypyrimidine tract (PPT) is essential for accurate 3' splice site identification. Progressive deletions within the PPT have been shown to impair correct lariat formation, spliceosome assembly, and subsequent splicing. In addition, the characteristics of the PPT can modulate 3' splice site selection by promoting the use of alternative branch sites [16,32,33].

Furthermore, there are several reports indicating that aberrant splicing can be associated with a missense or a silent change in an exon, rather than with mutations in introns. These findings suggest that such exonic mutations are capable of altering the secondary structure of the pre-mRNA, thereby interfering with the correct recognition of splicing regulatory elements [23,29,34].

3. *CSN1S1* and *CSN1S2* Gene Structure

Caseins account for about 80% of ruminant milk proteins. Moreover, they have a high nutritional value due to the contents in common and essential amino acids [35].

The α s-caseins are the most hydrophilic of all caseins. The primary distinction between the α s1-casein and α s2-casein families is predicated on their divergent amino acid sequences.

The encoding genes, *CSN1S1* and *CSN1S2*, are known to influence both the qualitative and quantitative properties of milk in major ruminant and non-ruminant species of zootechnical interest. These genes are annotated in nearly all species and are particularly well characterized in domestic ruminants (cattle, *Bos taurus*; river buffalo, *Bubalus bubalis*; sheep, *Ovis aries*; and goat, *Capra hircus*) and some camelids (dromedary, *Camelus dromedarius*; bactrian camel, *Camelus bactrianus*; llama, *Lama glama*). Both genes are located within a gene cluster spanning approximately 250 kb (kilobase), which also includes the β -casein and κ -casein genes (*CSN2* and *CSN3*, respectively). This cluster is located on chromosome 6 in *C. hircus*, *O. aries*, and *B. taurus*; on chromosome 7 in buffalo (*B. bubalis*) [36–40]; and on chromosome 2 in *C. dromedarius*, *C. bactrianus*, and *L. glama* [41], in which it spans a region

of about 190 kb. An exception to this chromosomal organization is found in mammals such as donkeys (*Equus asinus*), horses (*Equus caballus*), rabbits (*Oryctolagus cuniculus*), and rodents (*Rodentia*), which possess an additional copy of the *CSN1S2* gene. This indicates a recent paralogous gene duplication event [4,41–46].

It has been hypothesized that the *CSN1S1*, *CSN2*, and *CSN1S2* genes evolved from a primordial gene, likely through both intra and intergenic duplications of an ancestral gene consisting of four exons [47–50]. In contrast to *CSN2*, the *CSN1S1* and *CSN1S2* genes have undergone extensive exon duplication, exon skipping, and exon conversion events. These processes have resulted in an extremely fragmented gene architecture characterized by numerous small exons within the coding region. Due to this close evolutionary relationship, the 5' non-coding region, the domains encoding the ubiquitous multiple phosphorylation sites, and the 3' non-coding region of the mRNAs of *CSN1S1* and *CSN1S2* exhibit high levels of sequence homology and are well-conserved [49,51–56].

CSN1S1 and *CSN1S2* exhibit a conserved gene organization among ruminants, with some differences in intron size. These variations are primarily associated to artiodactyl retrotransposons (such as short interspersed elements—SINEs—and long interspersed elements—LINEs) located within their introns and regulatory regions [39,41,55]. In dairy ruminants, the *CSN1S1* gene spans approximately 16.7 kb and is composed of 19 exons and 18 introns. The exons show significant size variability, ranging from a minimum of 24 base pairs (bp) (exons 5, 6, 7, 8, 10, 13, and 16) to a maximum of 385 bp (exon 19). The first exon, a 53 bp sequence, is entirely non-coding. The leader peptide and the first two amino acids of the mature protein are encoded by exon 2 (63 bp). The translation stop codon TGA is created by coupling the final di-nucleotide TG of exon 17 with the first nucleotide A of exon 18 [55–57].

A similar structure also characterizes the *CSN1S1* gene in equids (*Equidae*) and camelids (*Camelidae*). In particular, in *E. caballus* and *E. asinus*, it has 19 exons, in *C. dromedarius* 20 exons, and in the other *Camelidae* 21 exons, ranging in size from 18 bp (exon 14) to 387 bp (exon 21) [9,41,46].

Similarly, in dairy ruminants, the *CSN1S2* gene is composed of 18 exons, which vary in size from 21 bp (exon 4) to 267 bp (exon 18). The first exon (44 bp) is entirely non-coding. The highly conserved signal peptide, a 15 amino acid sequence (MKFFIFTCLLAVALA), is encoded by exon 2 (63 bp). The mature protein consists of 207 amino acids in *B. taurus* and *B. bubalis*, and 208 in *O. aries* and *C. hircus*. The translation stop codon, TAA, is located at nucleotides 10 to 12 of exon 17. For both *CSN1S1* and *CSN1S2*, all splice junctions conform to the consensus GT-AG rule [39,45,51,54–59].

This gene has also recently been characterized in other minor dairy species. In *Camelidae*, the entire gene spans 17 exons, with sizes ranging from 24 bp (exons 4, 8, 11, and 13) to 280 bp (exon 17), and it encodes a mature peptide of 187 amino acids. Exon 1 (48 bp) is non-coding, as are the first 12 bp of exon 2. The signal peptide, a 15 amino acid sequence, initiates translation with an ATG codon at the 13th nucleotide and is fully encoded by the subsequent 45 nucleotides of exon 2. The mature peptide consists of 187 amino acids, is encoded by the last six nucleotides of exon 2, and continues up to the first 9 bp of exon 16. The translation stop codon, TAA, starts at nucleotides 10 of exon 16 [9,60].

As previously reported in the *E. asinus* genome, two different *CSN1S2* genes, called *CSN1S2I* and *CSN1S2II*, have been found. The first is constituted by 19 exons, and it encodes for α 2-CN-I, a protein of 221 amino acids, while the latter, consisting of 16 exons, is likely the result of a gene duplication event. In contrast to humans (*Homo sapiens*), where both *CSN1S2* paralogs are non-functional, the *E. asinus* *CSN1S2II* gene is considered functional. This functionality suggests that it produces an mRNA that is subsequently translated into

an α 2-CN II protein. Similarly to *E. caballus*, both genes are closely associated and have been mapped to chromosome 3 [4,44,46].

4. Variants Involving Canonical Splice Sites at *CSN1S1* and *CSN1S2* Loci Affecting Pre-mRNA Splicing

Among the different protein fractions of milk, caseins α 1 and α 2 represent those that have an evolutionary conserved role and, therefore, are the most investigated. Given their functional importance, intricate structure, and remarkable variability, the genes encoding these proteins serve as powerful molecular models for evolutionary research. They also provide valuable insights into the genetic architecture of lesser-studied species and aid in elucidating the phylogenetic relationships among mammalian and domestic animal species [61].

4.1. Polymorphisms at *CSN1S1* Locus

The α 1-CN family constitutes one of the most important milk casein fractions in dairy livestock species and consists of major and minor components [62,63]. Its encoding gene (*CSN1S1*) shows the highest level of polymorphism among those encoding the remaining milk proteins. The number of alleles identified at this locus has increased significantly over the years. To date, at least 23 variants of the *CSN1S1* gene have been identified in goats (*C. hircus*). These variants are associated with qualitative and quantitative differences in α 1-casein content and have been clustered into four distinct categories based on their expression levels: ‘strong’ (A, A1, A2, A3, A0, B0, 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). Additionally, two new α 1-casein variants, a ‘null’ and an ‘F-like’ variant, have recently been detected in an indigenous Norwegian goat population [64–66].

In *B. taurus* the *CSN1S1* locus shows ten alleles (A, B, C, D, E, F, G, H, I, and J), which correspond to nine protein variants. For the G variant specifically, a considerable reduction in α 1-casein content has been reported, which consequently leads to lower milk coagulation times and reduced curd firmness. This highlights the direct link between this genetic variant, its effect on protein quantity, and the resulting technological properties of the milk [67–70].

The nomenclature for milk proteins is unified across the four main species of the genus *B.* considered in cattle milk protein studies: *B. taurus* (taurine cattle), *B. indicus* (zebu), *B. grunniens* (yak), and *B. javanicus* (banteng) [68,70].

In *O. aries*, nine variants of the α 1-casein have been identified, namely A, B, C, D, E, F, G, H, and I, although not all have been fully characterized. In particular, the *CSN1S1* D allele is associated with a lower casein content, a reduced casein index, a greater whey protein content, and poorer lactodynamographic parameters [54,71–76].

This locus is equally polymorphic in buffalo. To date, the known variants are A, B, B′, B′′, C, D, E, and F, among which B′ and B′′ are synonymous variants of B, while eight amino acid deletions characterize variants E and F compared with the others. In particular, the F variant corresponds to the previously named α 1-CN variant BRV [77,78].

The ancestral B variant is shared by both river and swamp buffalo. Alleles A, B′, B′′, C, E, and F have been found in river buffalo, whereas the D variant is exclusive to swamp buffalo [58,77–79].

CSN1S1 gene polymorphism has also been extensively examined in camelids and equids. In *C. dromedarius*, at least four variants (α 1-caseins A, B, C, and D) have been identified [41,80–84], while only two variants have been reported in *L. glama* [85]. At this locus, six alleles (A, A*, B, C, D, and E) have been characterized in *E. caballus* [86]. In

contrast, *E. asinus* has been less investigated but is known to possess a variant associated with an apparently null content of this protein fraction [87].

4.2. Variants Involving Canonical Consensus Splicing Regulatory Sequences at CSN1S1 Locus

During livestock evolution, the *CSN1S1* gene acquired many sequence variations that lead to altered gene expression. To date, at least 13 alleles (Table 1) originating from mutations that affect consensus splicing regulatory sequences have been identified at this locus.

Table 1. Alleles by alteration of consensus splicing regulatory sequences at *CSN1S1* and *CSN1S2* loci.

Locus	Species (Common Name)	Alleles	Constitutive Exon Skipped	PEPTIDE Deleted *	Involved Consensus Splicing Sequences	References
CSN1S1	<i>Bos taurus</i> (Cattle)	A; A1 H	Exon 4 Exon 8	¹⁴ EVLNENLLRFFVA ²⁶ ⁵¹ NQAMENIK ⁵⁸	Donor splice sites	[88–90]
	<i>Ovis aries</i> (Sheep)	H I E	Exon 8 Exon 7 Exon 10	⁵¹ DQAMEDAK ⁵⁸ ⁴³ DIGSESIE ⁵⁰ ⁷⁰ EIVPNSAE ⁷⁷	Donor splice sites	[72,73,91]
	<i>Bubalus bubalis</i> (Water buffalo)	E (Bbt) F (Bbt)	Exon 6	³⁵ EKVNELST ⁴²	Donor splice site	[77,78]
	<i>Capra hircus</i> (Goat)	D G	Exon 9 Exon 4	⁵⁹ QMKAGSSSSSE ⁶⁹ ¹⁴ EVLNENLLRFFVA ²⁶	Donor splice sites	[92,93]
	<i>Camelus dromedarius</i> (Dromedary)	A; C; D	Exon 18	¹⁵⁵ EQAYFHLE ¹⁶²	Branch sequence and polypyrimidine tract	[80,81]
CSN1S2	<i>Bos taurus</i> (Cattle)	D	Exon 8	⁵¹ EYSIGSSSE ⁵⁹	Donor splice site	[94]
	<i>Bubalus bubalis</i> (Water buffalo)	B; B1; B2	Exon 7	⁴² EVIRNANEE ⁵⁰	Donor splice site	[39]
	<i>Capra hircus</i> (Goat)	D	Exon 11	⁸⁴ NEINQFYQKFPQYLQ YPYQGPIVLNPWDQVK RNAGPFTPTV ¹²⁴	Donor splice site	[95,96]
	<i>Equus asinus</i> (Donkey)	Unnamed	First 15 nucleotides of exon 17	¹⁷⁶ NKINQ ¹⁸⁰	Acceptor splice site	[4]
	<i>Camelus dromedarius</i> (Dromedary)	Unnamed	Uncharacterized	Uncharacterized		[83]

* aa number relative to mature proteins.

4.2.1. *Bos taurus*

The *CSN1S1* A variant of *B. taurus* lacks 13 amino acids (¹⁴EVLNENLLRFFVA²⁶) if compared with the B variant. This deletion stems from the corresponding absence of 39 bp (the entirety of exon 4) in the mRNA. This exon skipping event is directly correlated with a single point mutation (T > A transversion) at position +6 in the splice donor sequence distal to exon 4. This specific mutation impairs the normal splicing process of the A allele's α s1-casein pre-mRNA, leading to the skipping of the upstream exon [77,88]. The authors postulate that the A allele-specific mutation at position +6 disrupts the perfect complementarity between the intron-4 splice donor signal (positions +1 to +8) and U1-snRNP, thus preventing the U1-snRNP from compensating for a rather weak upstream splice acceptor sequence of exon 4, which is crucial for facilitating the initial binding of the U2AF65 protein to the polypyrimidine tract. This mechanism effectively explains why the A variant's mRNA lacks exon 4.

Furthermore, the analysis of the *CSN1S1* gene in a New Zealand herd producing A variant protein (A1) revealed another mutation, a single base deletion (Adenine) in position 4 of the 5' splice sequence of the same intron. However, the net result is identical, namely, exon 4 skipping during mRNA processing. It was hypothesized that, due to the deletion of

the 13 amino acid residues, the A variant lost a major chymosin cleavage site and had a less hydrophobic nature than its other counterparts in the α s1-casein family. Therefore, the physico-chemical properties may be different from other α s1-casein variants [89].

A similar molecular event was also presumed for a third allele at the bovine *CSN1S1* locus named H, which is characterized by the skipping of exon 8 encoding the ⁵¹NQAMENIK⁵⁸ peptide [90]. Bovine *CSN1S1* A and H alleles are not common. Variant A was reported only in Holstein Friesian, Red Holstein, and German Red breeds with an allele frequency between 0.01 and 0.001, while *CSN1S1* H was present only in German Friesian (0.002) and African Kuri cattle (0.04) [90,97–99].

4.2.2. *Ovis aries*

Similarly to the bovine species, the ovine *CSN1S1* locus shows three alleles characterized by mutations in canonical splicing sites *CSN1S1* E, H, and I.

The *CSN1S1* E variant lacks the amino acids 70 to 77 (EIVPNSAE), corresponding to the entirety of exon 10. This deletion also causes the loss of four phosphoserine (SerP) residues located at sites 64, 66, 68, and 75. The reduced content of α s1-casein E could therefore have a detrimental effect on the cheese-making properties of the milk and its mineral carrier activity. However, no further molecular genetic analyses are available [76,91,100].

As in the *B. taurus* *CSN1S1* H allele, the skipping of exon 8 also characterizes the *CSN1S1* H allele in the ovine species. It has been demonstrated that the deletion of four nucleotides (the last three of exon 8 and the first one of intron 8), replaced by a 13-nucleotide insertion in the DNA sequence, is the mutation underlying skipping of exon 8, leading to variant H production. This event inactivates the splice donor sequence distal to exon 8, leading to its skipping during the serial splicing reactions of the ovine α s1-casein pre-mRNA. Consequently, the mature protein lacks the peptide sequence ⁵¹DQAMEDAK⁵⁸, which is normally encoded by exon 8 [72].

Finally, the sequencing of *CSN1S1* cDNA and the mature protein led to the identification and characterization of the *CSN1S1* I allele, carrying the constitutive skipping of exon 7. This allelic aberration is correlated with a sequence difference (T > A) in the 5'-splice donor sequence of intron 7, which disrupts the crucial base pairing with U1-snRNA. As a result, the mature mRNA sequence of ovine *CSN1S1* I is 24 bp shorter than the wild-type coding sequence. This deletion leads to the absence of eight amino acids (⁴³DIGSESIE⁵⁰) in the mature protein, including two phosphorylated serine residues (p.SerP46 and p.SerP48) [73]. Beyond the loss of these two phosphoserine residues, the deleted sequence also contains three acidic residues (one Asp and two Glu). These residues are critical for forming electrostatic bonds with calcium ions, which contributes to higher micellar size and improved curd consistency. Consequently, this variant has a reported negative impact on all milk technological properties, with a significant effect on protein yield and content [73,74].

Furthermore, it was observed that the *CSN1S1* I allele is expressed at the same level as the full-length variants, while the deleted ovine *CSN1S1* E and H alleles show a lower expression [73,91]. Allele frequencies of the E, H, and I α s1-CN variants are generally low and occur only in some breeds. α s1-CN H seems to be characteristic for East Friesian dairy sheep (allele frequency from 0.08 to 0.1869), whereas α s1-CN I (allele frequency 0.029) and E (0.003) were identified only in the Gray Horned Heath and Leccese breeds, respectively [72–74,101,102].

4.2.3. *Capra hircus*

In *C. hircus*, two protein variants are known to originate from mutations at canonical splice sites at the *CSN1S1* locus. A deletion of 11 amino acids (⁵⁹QMKAGSSSSSE⁶⁹) characterizes the variant named *CSN1S1* D. This protein sequence is encoded by exon 9, which

comprises the major phosphorylation site of the protein [103]. Splice site mutations are, according to some reports, the cause of this deletion; they would, in fact, cause the alteration of RNA processing and the exclusion of exon 9 in the mature RNA assembly [92,103].

Similarly, the protein variant α s1-CN G lacks 13 amino acids (¹⁴EVLNENLLRFVVA²⁶) corresponding to exon 4. In this case, the skipping is caused by a G > A transition at the first nucleotide of the 5' splice site consensus sequence at the beginning of intron 4 [104].

Both *CSN1S1* G and D alleles are associated with low α s1-casein content in milk [93,104]. Unfortunately, to date, there are few studies evaluating the frequency of these alleles in different goat breeds/genetic types. The only report is related to the *CSN1S1* D allele, the frequency of which in a large Alpine herd of Moissac was 0.025. It was also observed with a low frequency in the Saanen breed [92].

4.2.4. *Bubalus bubalis*

In *B. bubalis*, to date, two alleles of the *CSN1S1* locus characterized by alteration of a canonical splice site are known: variants E and F. The first may directly originate from variant A through constitutive exon 6 skipping (involving 24 nucleotides), while variant F could originate from the B allele, as it shares the same amino acid substitutions at positions 31 and 44 but exhibits the same skipping phenomenon as the E allele. Based on exon skipping alone, both alleles are commonly grouped into a single allele, called *CSN1S1* Bbt or BRV [78]. This exon skipping event triggers the synthesis of a defective protein lacking eight amino acids (EKVNELST) between positions 35 and 42 of the mature protein. The molecular weight of the deleted mutant is 919 Da less than the wild type. The molecular event responsible for exon 6 skipping during mRNA maturation in *CSN1S1* Bbt alleles has been identified as a G to C substitution at the first position of intron 6. This mutation inactivates the donor splice site, leading to the observed exon skipping [77,105].

Stepic et al. [106] observed that *CSN1S1* Bbt was associated with higher protein and casein content in milk, but its effect on milk composition and its technological properties requires further investigation.

The *CSN1S1* Bbt allele has a frequency of 0.18 in Romanian buffaloes [77] and 0.23 in a Serbian buffalo population [106]; a similar frequency is also thought to occur in animal populations from Venezuela, Bulgaria, Poland, and Canada. In contrast, the deleted α s1-casein variant did not occur in a Mediterranean Italian River Buffalo breed reared in Italy [105,107].

4.2.5. *Camelus dromedarius*

The *CSN1S1* gene of *C. dromedarius* consists of 20 exons, unlike in other camelids, in which it is divided into 21 exons. The main difference arises from exon 20 (44 bp), which partially codes for the termination stop codon (exon 19 5'-TG . . . A-3' exon 20). The cause of this difference lies in a mutation that occurred exclusively at the donor splice site of the dromedary sequence, which alters the correct identification of the splicing sites and results in the skipping of the exon. The correct reading frame is subsequently restored by the next exon, which also begins with an adenine, thereby re-establishing the termination stop codon. In other camelids, both exon 20 and the corresponding splicing sites are conserved [9,41]. Therefore, it should not be considered an intra-species variant, but rather a constitutive event of *C. dromedarius*.

In addition, the presence of genomic variants at the *CSN1S1* locus, originating from mutations in consensus splicing regulatory sequences, has been hypothesized for dromedary camels. In particular, three alleles, *CSN1S1* A, C and D, would appear to be characterized by the absence of exon 18 (exon 16 using bovine exon numbering) at the mRNA level (encoding for the peptide ¹⁵⁵EQAYFHLE¹⁶²) [80,81] compared with the *CSN1S1* B allele found

by Kappeler et al. [108], with a consequent difference in the length of the corresponding protein (207 amino acids vs. 215). However, the sequence of exon 18 is present at the DNA level in all three alleles. Therefore, the protein variant could originate by the skipping out of the exon during mRNA processing, probably as a consequence of a not well characterized mutation at the DNA level altering the pre-mRNA spliceosome machinery. It is assumed that the mutational event underlying this variant is an 11 bp insertion (ATTGAATAAAA) located within intron 17, specifically positioned between the branch point sequence and the polypyrimidine tract immediately upstream of exon 18. The change in the α 1-CN peptide pattern due to the absence of this exon may modify the IgE-binding epitopes and alter the availability of bioactive peptides [81].

Furthermore, the *CSN1S1* C allele is characterized with respect to the remaining alleles at this locus by a non-synonymous single-nucleotide polymorphism G > T occurring at exon 5, leading to the amino acid substitution p.Glu30 > Asp in the mature protein and responsibility for Isoelectric Point change (3.22 vs. 2.77). On the other hand, no further information is reported for the D allele [80,81]. Therefore, it is to be considered that C and D are synonymous variants of the A.

The variant missing exon 18 seems to be the usual form within *C. dromedaries* *CSN1S1* mRNA and α 1-CN protein. In fact, the frequencies of the alleles without this exon have been found to be higher, with dominance of *CSN1S1* A (frequency ranging from 1.00 to 0.79) and a lower frequency of the C (0.00–0.28) and D (0.020–0.150) alleles [80,81,84].

4.3. Polymorphisms at *CSN1S2* Locus

The *CSN1S2* locus was also found to be particularly polymorphic in the four main dairy ruminant species. To date, in *O. aries*, seven alleles have been characterized (*CSN1S2* A, B, C, D, E, F, and G) [109], while at least eight alleles have been identified to date in buffalo: *CSN1S2* A, B, B1, B2, C, D, E, and F [39,110].

The most remarkable polymorphism in this locus has been found in *C. hircus*. Currently, several alleles associated with different α 2-CN expression levels have been characterized. Eight alleles of the *CSN1S2* gene are known in *C. hircus*, and they are associated with three different quantities of α 2-CN in milk. These variants are clustered as follows: normal (*CSN1S2* A, B, C, E, F, and G), corresponding to ~2.5 g/L of α 2-casein per allele; intermediate (D), corresponding to ~1.5 g/L per allele); or null (0), resulting in no detectable amount of this casein [95,111–114]. Four new non-defective α 2-CN variants (H, I, J, and K) were detected in domestic breeds and wild goat species reared in Sudan [115] and eight novel variants (L, M, N, O, P, Q, R, and S) were found in Indian goats [116].

To date, the studies carried on *B. taurus* have identified only five variants (A, B, C, E, and D) [62,69]. Among these, the B and C alleles seem to be specific for the *B. indicus* and *B. grunniens*, respectively [117].

The high level of polymorphism at the *CSN1S2* locus is not considered unusual. Studies of the α 2-casein have revealed extensive polymorphisms at both the protein and DNA levels across various species. Indeed, the characterization of the *CSN1S2* gene in Old World camels and the *CSN1S2I* and *II* genes in *E. asinus* and *E. caballus* has also revealed interesting genetic variations. In horses, a total of six mutations leading to eight distinct putative protein isoforms (*CSN1S2I* A, B, C, D1, D2, E1, E2, F) were identified. In particular, the horse *CSN1S2I* B variant is a α 2-CN short form as it is characterized by the presence of a 1.3 kb in-frame deletion spanning two coding exons corresponding to 17 amino acid residues [86,118,119].

The *CSN1S2* I variant of *E. asinus* has seven SNPs at the mRNA level, five of which led to amino acid changes, while at the *CSN1S2II* locus, nine SNPs were observed, seven of which are non-synonymous [4,46].

Currently, the molecular mutations that cause most of the above-mentioned *CSN1S1* and *CSN1S2* alleles are known, from single-nucleotide substitutions/deletions to large insertions/deletions.

It has been observed that the presence of specific genomic variants affecting the consensus splicing regulatory sequences in these two genes can modify the splicing process. This leads to partial or complete exon loss or intron gain in the mature mRNA, which ultimately alters the corresponding protein-coding sequence.

4.4. Variants Involving Consensus Splicing Regulatory Sequences at *CSN1S2* Locus

Although less frequently, polymorphisms that affect consensus splicing regulatory sequences are also documented at the *CSN1S2* locus in several livestock species of interest. This may be a consequence of the more limited number of studies conducted at this locus. In fact, to date, these variants have been described only in bovine, caprine, bubaline, asinine, and camelid species (Table 1).

4.4.1. *Bos taurus*

The deletion of the nine amino acid residues, ⁵¹EYSIGSSSE⁵⁹, characterize the α 2-CN D variant in *B. taurus*. The mutational event leading to the *CSN1S2* D allele is the transversion c.221G > T, located at the last nucleotide of exon 8, corresponding to the 5' consensus splicing site. This is considered the cause of the altered splicing of the primary transcript of α 2-CN D. Six taurine breeds (Angler, German Yellow, Hinterwälder, Limpurger, Hungarian Grey Steppe, and Vorderwälder) carry this allele, albeit at low frequency (ranging from 0.02 to 0.10), while it was not reported in indicine breeds [69,94].

4.4.2. *Capra hircus*

The *CSN1S2* D allele shows a 106 bp deletion, from the last 11 bp of exon 11 and the first 95 bp of the following intron. As a result of this deletion, exon 11 loses the canonical donor splice site. Moreover, the last undeleted nucleotide (A) of exon 11 (the first of codon 121) and the two nucleotides of the following intron (AT) could form a new codon corresponding to Asn121, followed by a new GT dinucleotide splicing donor site with lacking at least three codons coding for Pro122, Thr123, and Val124. This variant is 205 vs. 208 amino acids long [113]. However, a preliminary analysis of the transcripts produced by the *CSN1S2* D allele was unable to detect mRNA characterized by this event. Conversely, transcripts characterized by the skipping of the entirety of exon 11 (123 nt, coding for the peptide ⁸⁴NEINQFYQKFPQYLQYPYQGPIVLNPWDQVKRNAGPFTPTV¹²⁴) have been observed, which would consequently translate into a protein of only 167 amino acids [96].

Like most defective alleles arising from mutations affecting canonical splice sites, the goat *CSN1S2* D, apparently associated with decreased synthesis of α 2-CN in goat milk, is a rare allele. To date, its presence has been demonstrated in the “Argentata dell’Etna” (0.012) and “Girgentana” (0.006) breeds, as well as in an undefined genetic type belonging to a local population (frequency 0.019), all reared in southern Italy. It has also been found in Saanen (0.174) and Thai Native (0.143) goats raised in Thailand, and in Hungarian goats (0.005) [113,120–122]. In contrast, the allele appears to be absent in autochthonous goat breeds reared in Northern and Central Italy [66,123–126], as well as in populations from Turkey [127,128] and the Czech Republic [129].

4.4.3. *Bubalus bubalis*

The deleted allele *CSN1S2* B identified in *B. bubalis* carries an SNP (FM865620:g.773G > C) positioned at the donor splice site of exon 7 (27 bp). This mutation causes the complete skipping of exon 7 and the deletion of nine amino acids (⁴²EVIRNANEE⁵⁰) in mature protein. Consequently, the predicted mature bubaline α 2-CN B protein is 198 aa long

instead of 207 aa [110]. More recently, two other B-derived alleles have been described, named *CSN1S2* B1 and B2, characterized by the same mutation at the intron 7 splice donor site but differing for single polymorphisms (haplotypes), with B2 probably arising from interallelic recombination (single crossing) between the alleles D and B (or B1) [39]. An association study between the SNP g.773G > C and milk traits, including fatty acid composition, was carried out, but no significant associations were observed [39,130].

To date, the frequency of this mutation has been estimated only in Italian Mediterranean buffalo reared in Campania and Basilicata regions (Southern Italy), and it was found to be 0.17 [39,110].

4.4.4. *Equus asinus*

Mutation of splice junctions sometimes leads to activation of cryptic splice sites. In fact, it has been demonstrated that *E. asinus* carries a characteristic polymorphism (FM946022.1: c.375-1G > A) at the splice acceptor site of *CSN1S2I* exon 17. This point mutation results in allele-specific skipping of the first 15 nucleotides of this exon, which encode the peptide ¹⁷⁶NKINQ¹⁸⁰. Concurrently an in-frame cryptic splicing acceptor site (AACAAAATCAACCAG) is recognized. While the presence of the adenine, responsible for skipping the first 15 nucleotides of exon 17, is observed at a low frequency in *E. asinus* (0.2437), it appears to be a constitutive condition in all other species across different families within the *Perissodactyla* order. All these other *Perissodactyla* species are characterized by the presence of a second in-frame splice acceptor site, which enables recognition of the exon 17 deletion of the first 15 nucleotides during the splicing mechanism. Therefore, this case represents an intra-species allele-specific event, rather than a constitutive condition, as reported for *E. caballus*. In fact, it has been suggested that the G allele represents the ancestral condition of the *CSN1S2* gene. During the evolution of *Perissodactyla*, a G > A transition may have occurred, resulting in the fixation of adenine in all families of this order, except in *E. asinus* [4]. A similar evolutionary pattern was previously described in camelids at the same locus.

The frequency of adenine at the splice acceptor site of *CSN1S2I* exon 17 was found to be 0.2437 in Ragusana and Amiatina donkeys reared in different Italian regions [4].

4.4.5. *Camelus dromedarius*

Recently, also in *C. dromedarius*, a mutation affecting splice site at the *CSN1S2* locus has been identified: c.403-9_403-4delTTTCT. The splice site variant, which has a frequency of 0.323 in Tunisian dromedaries, is hypothesized to have a strong influence on protein function and stability. However, limited information has been reported regarding the specific functional consequences of this event [60,83].

5. Discussion and Future Perspectives

This review provides a comprehensive look at mutations affecting consensus splicing regulatory sequences in the *CSN1S1* and *CSN1S2* genes of dairy livestock species. The prevalent finding is that most mutations leading to constitutive exon skipping at these loci primarily impact the 5' donor splice sites. This pattern is consistent with human studies, in which point mutations at the 5' donor splice site are more common than those at the 3' acceptor site (62% vs. 26%). The functional consequence of mutations at donor sites is thought to be a significant reduction in the pairing between the splice site and the complementary sequence in the U1 small nuclear ribonucleoprotein particle (U1snRNP), a critical early step in the intricate process of mRNA splicing [131].

These alterations in pre-mRNA splicing patterns add to the already widely investigated effects of alternative splicing in various species. In goats, for example, specific

polymorphisms can introduce premature termination codons (PTCs), leading to the accumulation of a minor alternative splice form where an exon is often skipped. This can trigger nonsense-mediated decay (NMD), profoundly impacting the expression levels of the encoded α s1 and α s2 caseins [7,95,132].

A growing number of health benefits and a diverse range of potential activities are attributed to milk proteins and the bioactive peptides encoded within their sequences. Caseins, particularly α s1-CN and α s2-CN, are known to be significant sources of these peptides [15,133]. Consequently, it is plausible that the truncated proteins resulting from constitutive exon skipping—a direct consequence of mutations in splice sites—could lead to the emergence of new allergenic epitopes or the alteration of existing IgE-binding epitopes. This would, in turn, affect the availability of valuable bioactive peptides. This hypothesis is supported by similar proposals for various protein isoforms with rearranged amino acid sequences that arise from normal alternative splicing events characteristic of casein-coding loci in different dairy livestock species [134]. Likewise, allele-dependent alternative splicing events could lead to similar outcomes. For instance, skipping of exon 18 of the A, C, and D alleles at the *CSN1S1* locus in the *C. dromedarius*, or of the first 15 nucleotides of exon 17 encoding the pentapeptide $^{176}\text{NKINQ}^{180}$ in the *E. asinus* α s2-CN variant, is hypothesized to change the casein peptide pattern and affect IgE-binding epitopes, thereby altering the availability of bioactive peptides [4,81].

In general, alternative splicing results in protein isoforms that typically characterize the protein pool from a single gene to a lesser extent, regardless of genetic variability. Conversely, with allele-specific splicing, the altered protein produced represents the primary translated product and possesses a unique amino acid sequence. Because this is a direct consequence of a precise mutational event, it becomes possible to conduct DNA-level selection for or against a specific variant to improve the functional quality of milk and its derived products. The loss of an entire exon or a portion of it can drastically alter the protein's three-dimensional structure, stability, and post-translational modifications (e.g., phosphorylation sites), leading to a cascade of effects on its biological function.

Future research should focus on a multidisciplinary approach to fully understand the implications of these findings. Identifying novel allele-specific splicing events across diverse dairy livestock species is a critical first step. Advanced omics technologies will be instrumental in this endeavor. Long-read transcriptomics (e.g., PacBio, Oxford Nanopore) is particularly valuable as it allows for the sequencing of full-length transcripts without the need for assembly, providing an accurate and complete profile of all splice isoforms from a single allele. Allele-resolved RNA sequencing will be crucial for distinguishing between the transcripts produced by different alleles within a heterozygous individual, offering an unprecedented level of detail on the regulatory mechanisms at these loci.

Translating genetic findings into practical innovations requires a deeper understanding of the functional consequences. Future studies should integrate proteomics, using techniques like mass spectrometry, to confirm that the identified mRNA isoforms are translated into functional proteins and to characterize their precise amino acid sequences and post-translational modifications. Furthermore, in vitro and in vivo functional studies will be essential to evaluate the nutritional, allergenic, and technological implications of these modified proteins. For example, cellular models could be used to test the immunogenicity of milk peptides from different alleles, while animal studies could assess the allergenic potential of milk from specific genotypic variants.

A deeper understanding of allele-specific splicing may enable precision breeding strategies aimed at producing milk with improved functional properties, reduced allergenicity, or enhanced technological performance for cheesemaking. The ability to select animals based on specific genotypic markers linked to beneficial splicing outcomes offers

a powerful tool for modern livestock management. Such advances could support the development of differentiated dairy products tailored to specific consumer needs, such as low-allergenicity milk or milk with a higher content of specific bioactive peptides. This approach not only adds value to the dairy supply chain but also contributes to more sustainable and economically viable production systems.

Author Contributions: Conceptualization, G.C., S.A., and A.P.; writing—original draft preparation, G.C., A.F., E.D., S.A., F.C., and A.P.; writing—review and editing, G.C., A.F., E.D., S.A., F.C., and A.P.; supervision, G.C., S.A., and A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Abbreviations

The following abbreviations are used in this manuscript:

pre-mRNA	precursor messenger RNA
snRNPs	small nuclear ribonucleoprotein particles
SINEs	short interspersed elements
LINEs	long interspersed elements
bp	base pairs
PTCs	premature termination codons

References

- Jiang, W.; Chen, L. Alternative splicing: Human disease and quantitative analysis from high-throughput sequencing. *Comput. Struct. Biotechnol. J.* **2020**, *24*, 183–195. [\[CrossRef\]](#)
- Gamazon, E.R.; Stranger, B.E. Genomics of alternative splicing: Evolution, development and pathophysiology. *Hum. Genet.* **2014**, *133*, 679–687. [\[CrossRef\]](#)
- Sibley, C.R.; Blazquez, L.; Ule, J. Lessons from non-canonical splicing. *Nat. Rev. Genet.* **2016**, *17*, 407–421. [\[CrossRef\]](#)
- Cosenza, G.; Pauciullo, A. A Comprehensive Analysis of CSN1S2 I and II Transcripts Reveals Significant Genetic Diversity and Allele-Specific Exon Skipping in Ragusana and Amiatina Donkeys. *Animals* **2024**, *14*, 2918. [\[CrossRef\]](#)
- López-Bigas, N.; Audit, B.; Ouzounis, C.; Parra, G.; Guigó, R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* **2005**, *28*, 1900–1903. [\[CrossRef\]](#)
- Raponi, M.; Baralle, D. Alternative splicing: Good and bad effects of translationally silent substitutions. *FEBS J.* **2010**, *277*, 836–840. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Johnsen, L.B.; Rasmussen, L.K.; Petersen, T.E.; Berglund, L. Characterization of Three Types of Human α (S1)-Casein mRNA Transcripts. *Biochem. J.* **1995**, *309*, 237–242. [\[CrossRef\]](#)
- Pauciullo, A.; Erhardt, G. Molecular Characterization of the Llamas (Lama Glama) Casein Cluster Genes Transcripts (CSN1S1, CSN2, CSN1S2, CSN3) and Regulatory Regions. *PLoS ONE* **2015**, *10*, e0124963. [\[CrossRef\]](#)
- Wang, Y.; Liu, J.; Huang, B.; Xu, Y.M.; Li, J.; Huang, L.F.; Lin, J.; Zhang, J.; Min, Q.H.; Yang, W.M.; et al. Mechanism of Alternative Splicing and its Regulation. *Biomed. Rep.* **2015**, *3*, 152–158. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kadri, N.K.; Mapel, X.M.; Pausch, H. The Intronic Branch Point Sequence Is under Strong Evolutionary Constraint in the Bovine and Human Genome. *Commun. Biol.* **2021**, *4*, 1206. [\[CrossRef\]](#)
- Riolo, G.; Cantara, S.; Ricci, C. What's Wrong in a Jump? Prediction and Validation of Splice Site Variants. *Methods Protoc.* **2021**, *4*, 62. [\[CrossRef\]](#)
- Marasco, L.E.; Kornblihtt, A.R. The Physiology of Alternative Splicing. *Nat. Rev. Mol. Cell Biol.* **2023**, *24*, 242–254. [\[CrossRef\]](#) [\[PubMed\]](#)
- Anna, A.; Monika, G. Splicing Mutations in Human Genetic Disorders: Examples, Detection, and Confirmation. *J. Appl. Genet.* **2018**, *59*, 253–268. [\[CrossRef\]](#)
- Burset, M.; Seledtsov, I.A.; Solovyev, V.V. Analysis of Canonical and Non-Canonical Splice Sites in Mammalian Genomes. *Nucleic Acids Res.* **2000**, *28*, 4364–4375. [\[CrossRef\]](#)

16. Sickmier, E.A.; Frato, K.E.; Shen, H.; Paranawithana, S.R.; Green, M.R.; Kielkopf, C.L. Structural Basis for Polypyrimidine Tract Recognition by the Essential Pre-mRNA Splicing Factor U2AF65. *Mol. Cell* **2006**, *23*, 49–59. [\[CrossRef\]](#)
17. Welden, J.; Stamm, S. Pre-mRNA structures forming circular RNAs. *Biochim. Biophys. Acta (BBA) Gene Regul. Mech.* **2019**, *1862*, 194410. [\[CrossRef\]](#)
18. Zhang, P.; Zhang, X.O.; Jiang, T.; Cai, L.; Huang, X.; Liu, Q.; Li, D.; Lu, A.; Liu, Y.; Xue, W.; et al. Comprehensive identification of alternative back-splicing in human tissue transcriptomes. *Nucleic Acids Res.* **2020**, *28*, 1779–1789. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Zhou, W.Y.; Cai, Z.R.; Liu, J.; Wang, D.S.; Ju, H.Q.; Xu, R.H. Circular RNA: Metabolism, functions and interactions with proteins. *Mol. Cancer* **2020**, *14*, 172. [\[CrossRef\]](#)
20. Pan, Q.; Shai, O.; Lee, L.J.; Frey, B.J.; Blencowe, B.J. Deep Surveying of Alternative Splicing Complexity in the Human Transcriptome by High-Throughput Sequencing. *Nat. Genet.* **2008**, *40*, 1413–1415. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Keren, H.; Lev-Maor, G.; Ast, G. Alternative Splicing and Evolution: Diversification, Exon Definition and Function. *Nat. Rev. Genet.* **2010**, *11*, 345–355. [\[CrossRef\]](#)
22. Wang, Z.; Burge, C.B. Splicing Regulation: From a Parts List of Regulatory Elements to an Integrated Splicing Code. *RNA* **2008**, *14*, 802–813. [\[CrossRef\]](#)
23. Lenasi, T.; Peterlin, B.M.; Dovc, P. Distal Regulation of Alternative Splicing by Splicing Enhancer in Equine β -Casein Intron 1. *RNA* **2006**, *12*, 498. [\[CrossRef\]](#)
24. O'Neill, J.P.; Rogan, P.K.; Cariello, N.; Nicklas, J.A. Mutations that alter RNA splicing of the human HPRT gene: A review of the spectrum. *Mutat. Res.* **1998**, *411*, 179–214. [\[CrossRef\]](#)
25. Pandey, S.; Shreya, S.; Jain, B.P. Post-Transcriptional Gene Regulation: An Overview. *Post Transcr. Gene Regul. Hum. Dis.* **2022**, *32*, 33–53. [\[CrossRef\]](#)
26. Kornblihtt, A.R.; Schor, I.E.; Alló, M.; Dujardin, G.; Petrillo, E.; Muñoz, M.J. Alternative Splicing: A Pivotal Step between Eukaryotic Transcription and Translation. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 153–165. [\[CrossRef\]](#)
27. Park, E.; Pan, Z.; Zhang, Z.; Lin, L.; Xing, Y. The Expanding Landscape of Alternative Splicing Variation in Human Populations. *Am. J. Hum. Genet.* **2018**, *102*, 11–26. [\[CrossRef\]](#)
28. Krawczak, M.; Thomas, N.S.T.; Hundrieser, B.; Mort, M.; Wittig, M.; Hampe, J.; Cooper, D.N. Single Base-Pair Substitutions in Exon-Intron Junctions of Human Genes: Nature, Distribution, and Consequences for mRNA Splicing. *Hum. Mutat.* **2007**, *28*, 150–158. [\[CrossRef\]](#)
29. Nembaware, V.; Wolfe, K.H.; Bettoni, F.; Kelso, J.; Seoighe, C. Allele-Specific Transcript Isoforms in Human. *FEBS Lett.* **2004**, *577*, 233–238. [\[CrossRef\]](#)
30. Nembaware, V.; Lupindo, B.; Schouest, K.; Spillane, C.; Scheffler, K.; Seoighe, C. Genome—Wide Survey of Allele-Specific Splicing in Humans. *BMC Genom.* **2008**, *9*, 265. [\[CrossRef\]](#)
31. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. ACMG Laboratory Quality Assurance Committee, Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. [\[CrossRef\]](#)
32. Ritchie, D.B.; Schellenberg, M.J.; MacMillan, A.M. Spliceosome Structure: Piece by Piece. *Biochim. Biophys. Acta Gene Regul. Mech.* **2009**, *1789*, 624–633. [\[CrossRef\]](#)
33. Coolidge, C.J.; Seely, R.J.; Patton, J.G. Functional Analysis of the Polypyrimidine Tract in Pre-mRNA Splicing. *Nucleic Acids Res.* **1997**, *25*, 888–895. [\[CrossRef\]](#)
34. Cosenza, G.; Martin, P.; Garro, G.; Gallo, D.; Auzino, B.; Ciampolini, R.; Pauciullo, A. A Novel Allelic Donkey β -Lactoglobulin I Protein Isoform Generated by a Non-AUG Translation Initiation Codon Is Associated with a Nonsynonymous SNP. *J. Dairy Sci.* **2023**, *106*, 4158–4170. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Comunian, T.A.; Favaro-Trindade, C.S. Microencapsulation Using Biopolymers as an Alternative to Produce Food Enhanced with Phytosterols and Omega-3 Fatty Acids: A Review. *Food Hydrocoll.* **2016**, *61*, 442–457. [\[CrossRef\]](#)
36. Threadgill, D.W.; Womack, J.E. Genomic Analysis of the Major Bovine Milk Protein Genes. *Nucleic Acids Res.* **1990**, *18*, 6935–6942. [\[CrossRef\]](#)
37. Leveziel, H.; Metenier, L.; Guerin, G.; Cullen, P.; Provot, C.; Bertaud, M.; Mercier, J.C. Restriction Fragment Length Polymorphism of Ovine Casein Genes: Close Linkage between the Alpha S1-, Alpha S2-, Beta- and Kappa-Casein Loci. *Anim. Genet.* **1991**, *22*, 1–10. [\[CrossRef\]](#)
38. Bevilacqua, C.; Ferranti, P.; Garro, G.; Veltri, C.; Lagonigro, R.; Leroux, C.; Pietrolà, E.; Addeo, F.; Pilla, F.; Chianese, L.; et al. Interallelic Recombination is Probably Responsible for the Occurrence of a New α s1-Casein Variant Found in the Goat Species. *Eur. J. Biochem.* **2002**, *269*, 1293–1303. [\[CrossRef\]](#) [\[PubMed\]](#)
39. 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. [\[CrossRef\]](#)

40. Rehman, S.U.; Hassan, F.U.; Luo, X.; Li, Z.; Liu, Q. Whole-Genome Sequencing and Characterization of Buffalo Genetic Resources: Recent Advances and Future Challenges. *Animals* **2021**, *11*, 904. [\[CrossRef\]](#)
41. Pauciullo, A.; Shuiep, E.T.; Ogah, M.D.; Cosenza, G.; Stasio, L.D.; Erhardt, G. Casein Gene Cluster in Camelids: Comparative Genome Analysis and New Findings on Haplotype Variability and Physical Mapping. *Front. Genet.* **2019**, *10*, 748. [\[CrossRef\]](#)
42. Stewart, A.F.; Bonsing, J.; Beattie, C.W.; Shah, F.; Willis, I.M.; Mackinlay, A.G. Complete Nucleotide Sequences of Bovine α (S2)- and β -Casein cDNAs: Comparisons with Related Sequences in Other Species. *Mol. Biol. Evol.* **1987**, *4*, 231–241. [\[CrossRef\]](#)
43. Ginger, M.R.; Piotte, C.P.; Otter, D.E.; Grigor, M.R. Identification, Characterisation and cDNA Cloning of Two Caseins from the Common Brushtail Possum (*Trichosurus vulpecula*). *Biochim. Biophys. Acta Gen. Subj.* **1999**, *1427*, 92–104. [\[CrossRef\]](#)
44. Cosenza, G.; Pauciullo, A.; Annunziata, A.L.; Rando, A.; Chianese, L.; Marletta, D.; Iannolino, G.; Nicodemo, D.; Di Berardino, D.; Ramunno, L. Identification and Characterization of the Donkey CSN1S2 I and II CDNAS. *Ital. J. Anim. Sci.* **2010**, *9*, 206–211. [\[CrossRef\]](#)
45. Rehman, S.U.; Feng, T.; Wu, S.; Luo, X.; Lei, A.; Luobu, B.; Hassan, F.-U.; Liu, Q. Comparative Genomics, Evolutionary and Gene Regulatory Regions Analysis of Casein Gene Family in Bubalus Bubalis. *Front. Genet.* **2021**, *12*, 662609. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Auzino, B.; Miranda, G.; Henry, C.; Krupova, Z.; Martini, M.; Salari, F.; Cosenza, G.; Ciampolini, R.; Martin, P. Top-Down Proteomics Based on LC-MS Combined with cDNA Sequencing to Characterize Multiple Proteoforms of Amiate Donkey Milk Proteins. *Food Res. Int.* **2022**, *160*, 111611. [\[CrossRef\]](#)
47. Gaye, P.; Gautron, J.P.; Mercier, J.C.; Hazé, G. Amino Terminal Sequences of the Precursors of Ovine Caseins. *Biochem. Biophys. Res. Commun.* **1977**, *79*, 903–911. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Hobbs, A.A.; Rosen, J.M. Sequence of Rat α - and γ -Casein mRNAs: Evolutionary Comparison of the Calcium-Dependent Rat Casein Multigene Family. *Nucleic Acids Res.* **1982**, *10*, 8079–8098. [\[CrossRef\]](#)
49. Jones, W.K.; Yu-Lee, L.Y.; Clift, S.M.; Brown, T.L.; Rosen, J.M. The Rat Casein Multigene Family. Fine Structure and Evolution of the Beta-Casein Gene. *J. Biol. Chem.* **1985**, *260*, 7042–7050. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Bonsing, J.; Mackinlay, A.G. Recent Studies on Nucleotide Sequences Encoding the Caseins. *J. Dairy Res.* **1987**, *54*, 447–461. [\[CrossRef\]](#)
51. Groenen, M.A.M.; Dijkhof, R.J.M.; Verstege, A.J.M.; van der Poel, J.J. The Complete Sequence of the Gene Encoding Bovine α s2-Casein. *Gene* **1993**, *123*, 187–193. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Mercier, J.C.; Vilotte, J.L. Structure and Function of Milk Protein Genes. *J. Dairy Sci.* **1993**, *76*, 3079–3098. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Passey, R.; Glenn, W.; Mackinlay, A. Exon Skipping in the Ovine α (S1)-Casein Gene. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* **1996**, *114*, 389–394. [\[CrossRef\]](#)
54. Rijnkels, M. Multispecies Comparison of the Casein Gene Loci and Evolution of Casein Gene Family. *J. Mammary Gland. Biol. Neoplasia* **2002**, *7*, 327–345. [\[CrossRef\]](#)
55. Ramunno, L.; Cosenza, G.; Rando, A.; Illario, R.; Gallo, D.; Di Berardino, D.; Masina, P. The Goat α s1-Casein Gene: Gene Structure and Promoter Analysis. *Gene* **2004**, *334*, 105–111. [\[CrossRef\]](#)
56. Calvo, J.H.; Dervishi, E.; Sarto, P.; González-Calvo, L.; Berzal-Herranz, B.; Molino, F.; Serrano, M.; Joy, M. Structural and Functional Characterisation of the α s1-Casein (CSN1S1) Gene and Association Studies with Milk Traits in Assaf Sheep Breed. *Livest. Sci.* **2013**, *157*, 1–8. [\[CrossRef\]](#)
57. Koczan, D.; Hobom, G.; Seyfert, H.M. Genomic Organisation of the Bovine Alpha-S1 Casein Gene. *Nucleic Acids Res.* **1991**, *19*, 5591–5596. [\[CrossRef\]](#)
58. 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*, 654–660. [\[CrossRef\]](#)
59. Stafuzza, N.; Naressi, B.; Amaral-Trusty, M. Sequence Analysis of the Alpha-S1 Casein Gene in Murrah Buffalo. In Proceedings of the 52a Reunião Anual da Sociedade Brasileira de Zootecnia, Belo Horizonte, Brazil, 19–23 July 2015.
60. Pauciullo, A.; Versace, C.; Gaspa, G.; Letaief, N.; Bedhief-Romdhani, S.; Fulgione, A.; Cosenza, G. Sequencing and Characterization of α s2-Casein Gene (CSN1S2) in the Old-World Camels Have Proven Genetic Variations Useful for the Understanding of Species Diversification. *Animals* **2023**, *13*, 2805. [\[CrossRef\]](#)
61. Cosenza, G.; Fulgione, A.; Albarella, S.; Ciotola, F.; Peretti, V.; Gallo, D.; Pauciullo, A. Identification and Validation of Genus/Species-Specific Short InDels in Dairy Ruminants. *BMC Vet. Res.* **2025**, *21*, 215. [\[CrossRef\]](#)
62. 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. *J. Dairy Sci.* **2004**, *87*, 1641–1674. [\[CrossRef\]](#)
63. Cosenza, G.; Mauriello, R.; Garro, G.; Auzino, B.; Iannaccone, M.; Costanzo, A.; Chianese, L.; Pauciullo, A. Casein Composition and Differential Translational Efficiency of Casein Transcripts in Donkey's Milk. *J. Dairy Res.* **2019**, *86*, 201–207. [\[CrossRef\]](#)
64. 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. [\[CrossRef\]](#)

65. Devold, T.G.; Nordbø, R.; Langsrud, T.; Svenning, C.; Brovold, M.J.; Sørensen, E.S.; Christensen, B.; Adnøy, T.; Vegarud, G.E. Extreme Frequencies of the α s1-Casein “Null” Variant in Milk from Norwegian Dairy Goats—Implications for Milk Composition, Micellar Size and Renneting Properties. *Dairy Sci. Technol.* **2010**, *91*, 39–51. [\[CrossRef\]](#)
66. Cosenza, G.; Albarella, S.; D’Anza, E.; Iannuzzi, A.; Selvaggi, M.; Pugliano, M.; Galli, T.; Saralli, G.; Ciotola, F.; Peretti, V. A New AS-PCR Method to Detect CSN201 Allele, Genotyping at Ca-Sensitive Caseins Loci and Milk Traits Association Studies in Autochthonous Lazio Goats. *Animals* **2023**, *13*, 239. [\[CrossRef\]](#)
67. 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*, 1735–1742. [\[CrossRef\]](#)
68. Caroli, A.M.; Chessa, S.; Erhardt, G.J. Invited Review: Milk Protein Polymorphisms in Cattle: Effect on Animal Breeding and Human Nutrition. *J. Dairy Sci.* **2009**, *92*, 5335–5352. [\[CrossRef\]](#)
69. Gallinat, J.L.; Qanbari, S.; Drögemüller, C.; Pimentel, E.C.G.; Thaller, G.; Tetens, J. DNA-Based Identification of Novel Bovine Casein Gene Variants. *J. Dairy Sci.* **2013**, *96*, 699–709. [\[CrossRef\]](#) [\[PubMed\]](#)
70. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Ramunno, L.; Cosenza, G.; Rando, A.; Macciotta, N.P.P.; Pappalardo, M.; Masina, P. Identification of Carriers of the Welsh CASA1 Variant Using an Allele-Specific PCR Method. *Anim. Genet.* **1997**, *28*, 154–155.
72. Giambra, I.J.; Chianese, L.; Ferranti, P.; Erhardt, G. Short Communication: Molecular Genetic Characterization of Ovine α s1-Casein Allele H Caused by Alternative Splicing. *J. Dairy Sci.* **2010**, *93*, 792–795. [\[CrossRef\]](#)
73. Giambra, I.J.; Chianese, L.; Ferranti, P.; Erhardt, G. Genomics and Proteomics of Deleted Ovine CSN1S1*I. *Int. Dairy J.* **2010**, *20*, 195–202. [\[CrossRef\]](#)
74. Giambra, I.J.; Brandt, H.; Erhardt, G. Milk Protein Variants Are Highly Associated with Milk Performance Traits in East Friesian Dairy and Lacaune Sheep. *Small Rumin. Res.* **2014**, *121*, 382–394. [\[CrossRef\]](#)
75. 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.* **2014**, *94*, 3090–3099. [\[CrossRef\]](#)
76. De Pascale, S.; Caira, S.; Garro, G.; Mauriello, R.; Scaloni, A.; Cosenza, G.; Chianese, L. Proteomic Characterisation and Phylogenetic Derivation of Ovine α s1-CN B and α s1-CN G Genetic Variants. *Int. Dairy J.* **2022**, *131*, 105387. [\[CrossRef\]](#)
77. Balteanu, V.A.; Carsai, T.C.; Vlaic, A. Identification of an Intronic Regulatory Mutation at the Buffalo α s1-Casein Gene That Triggers the Skipping of Exon 6. *Mol. Biol. Rep.* **2013**, *40*, 4311–4316. [\[CrossRef\]](#)
78. 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*). *Pak. J. Zool.* **2021**, *53*, 1233–1242. [\[CrossRef\]](#)
79. Chianese, L.; Quarto, M.; Pizzolongo, F.; Calabrese, M.G.; Caira, S.; Mauriello, R.; De Pascale, S.; Addeo, F. Occurrence of Genetic Polymorphism at the α s1-Casein Locus in Mediterranean Water Buffalo Milk. *Int. Dairy J.* **2009**, *19*, 181–189. [\[CrossRef\]](#)
80. Shuieip, E.T.S.; Giambra, I.J.; El Zubeir, I.E.Y.M.; Erhardt, G. Biochemical and Molecular Characterization of Polymorphisms of α s1-Casein in Sudanese Camel (*Camelus dromedarius*) Milk. *Int. Dairy J.* **2013**, *28*, 88–93. [\[CrossRef\]](#)
81. Erhardt, G.; Shuieip, E.T.S.; Lisson, M.; Weimann, C.; Wang, Z.; El Zubeir, I.E.Y.M.; Pauciullo, A. Alpha S1-Casein Polymorphisms in Camel (*Camelus dromedarius*) and Descriptions of Biological Active Peptides and Allergenic Epitopes. *Trop. Anim. Health Prod.* **2016**, *48*, 879–887. [\[CrossRef\]](#)
82. Mehta, S.C.; Dahiya, S.S.; Jadhav, S.A.; Umrikar, U.D.; Sawane, M.P.; Pawar, V.D.; Deshmukh, R.S. Determination of Genetic Variability at α s1-Casein Gene in Indian Dromedary. *Indian J. Anim. Sci.* **2020**, *90*, 1617–1621. [\[CrossRef\]](#)
83. Mutery, A.A.; Rais, N.; Mohamed, W.K.E.; Abdelaziz, T. Genetic Diversity in Casein Gene Cluster in a Dromedary Camel (*C. dromedarius*) Population from the United Arab Emirates. *Genes* **2021**, *12*, 1417. [\[CrossRef\]](#)
84. Amandykova, M.; Dossybayev, K.; Mussayeva, A.; Bekmanov, B.; Saitou, N. Comparative Analysis of the Polymorphism of the Casein Genes in Camels Bred in Kazakhstan. *Diversity* **2022**, *14*, 285. [\[CrossRef\]](#)
85. Pauciullo, A.; Gauly, M.; Cosenza, G.; Wagner, H.; Erhardt, G. Lama Glama α s1-Casein: Identification of New Polymorphisms in the CSN1S1 Gene. *J. Dairy Sci.* **2017**, *100*, 1282–1289. [\[CrossRef\]](#)
86. Brinkmann, J.; Jagannathan, V.; Drögemüller, C.; Rieder, S.; Leeb, T.; Thaller, G.; Tetens, J. Genetic Variability of the Equine Casein Genes. *J. Dairy Sci.* **2016**, *99*, 5486–5497. [\[CrossRef\]](#)
87. Ramunno, L.; Chianese, L.; Di Gregorio, P.; Rando, A.; Marletta, D.; Mauriello, R.; Quarto, M.; Gallo, D.; Pauciullo, A.; Cosenza, G.; et al. Stato dell’arte sulla Caratterizzazione Proteomica e Genomica delle Proteine del Latte di Asina. In *Latte di Asina: Produzione, Caratteristiche e Gestione Dell’azienda Asinina*; Fondazione Iniziative Zooprofilattiche e Zootecniche: Brescia, Italy, 2011; pp. 183–191, ISBN 9788890441660.
88. Mohr, U.; Koczan, D.; Linder, D.; Hobom, G.; Erhardt, G. A Single Point Mutation Results in a Allele-Specific Exon Skipping in the Bovine α s1-Casein mRNA. *Gene* **1994**, *143*, 187–192. [\[CrossRef\]](#)

89. Wilkins, R.; Xie, T. Two Distinct Gene Mutations—One Milk Protein Polymorphism: The Example of the α 1-Casein A Variant. In *IDF Seminar*; International Dairy Federation: Palmerston North, New Zealand, 1997; pp. 330–333.
90. Mahé, M.F.; Miranda, G.; Queval, R.; Bado, A.; Zafindrajaona, P.S.; Grosclaude, F. Genetic Polymorphism of Milk Proteins in African *Bos taurus* and *Bos indicus* Populations. Characterization of Variants α 1-Cn H and κ -Cn J. *Genet. Sel. Evol.* **1999**, *31*, 239–253. [\[CrossRef\]](#)
91. Chianese, L.; Caira, S.; Garro, G.; Lilla, S.; Addeo, F. Primary Structure of Ovine Deleted Variant α 1-CN E. In Proceedings of the 5th International Symposium on the Challenge to Sheep and Goats Milk Sectors, Alghero, Italy, 18–20 April 2007; pp. 58–60.
92. Mahé, M.F.; Grosclaude, F. α 1-CnD, Another Allele Associated with a Decreased Synthesis Rate at the Caprine α 1-Casein Locus. *Genet. Sel. Evol.* **1989**, *21*, 127–129. [\[CrossRef\]](#)
93. Martin, P.; Leroux, C. Characterization of a Further Alpha-S1-Casein Variant Generated by Exon-Skipping. In Proceedings of the 24th International Society of Animal Genetics, Prague, Czech Republic, 23–29 July 1994; Abstract E43. p. 8.
94. Bouniol, C.; Printz, C.; Mercier, J.C. Bovine α 2-Casein D Is Generated by Exon VIII Skipping. *Gene* **1993**, *128*, 289–293. [\[CrossRef\]](#)
95. 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 Casein in of α 2 Casein in Goat Milk. *Anim. Genet.* **2001**, *32*, 19–26. [\[CrossRef\]](#)
96. Cosenza, G.; Gallo, D.; Riccio, R.; Di Berardino, D.; Bisogno, A.; Ramunno, L. Analisi Preliminari dei Trascritti di Alcuni Alleli al Locus CSN1S2 di Capra. In Proceedings of the XV Congresso Nazionale SIPAOC, Cagliari, Italy, 11–14 September 2002.
97. Grosclaude, F. Le Polymorphisme Génétique Des Principales Lactoprotéines Bovines. Relations Avec La Quantité, La Composition et les Aptitudes Fromagères du Lait. *INRAE Prod. Anim.* **1988**, *1*, 5–17. [\[CrossRef\]](#)
98. Erhardt, G. A New α 1-casein Allele in Bovine Milk and its Occurrence in Different Breeds. *Anim. Genet.* **1993**, *24*, 65–66. [\[CrossRef\]](#)
99. Formaggioni, P.; Summer, A.; Malacarne, M.; Mariani, P. Milk Protein Polymorphism: Detection and Diffusion of the Genetic Variants in Bos Genus. *Ann. Fac. Med. Vet. Univ. Parma* **1999**, *19*, 127–165.
100. Garro, G.; Caira, S.; Lilla, S.; Mauriello, R.; Chianese, L. Characterisation of the Heterogeneity of Ovine Deleted Variant α 1-Casein E by a Proteomic Approach. *Int. Dairy J.* **2019**, *89*, 53–59. [\[CrossRef\]](#)
101. Giambra, I.J.; Jäger, S.; Erhardt, G. Isoelectric Focusing Reveals Additional Casein Variants in German Sheep Breeds. *Small Rumin. Res.* **2010**, *90*, 11–17. [\[CrossRef\]](#)
102. Chianese, L.; Garro, G.; Mauriello, R.; Laezza, P.; Ferranti, P.; Addeo, F. Occurrence of Five α 1-Casein Variants in Ovine Milk. *J. Dairy Res.* **1996**, *63*, 49–59. [\[CrossRef\]](#)
103. Brignon, G.; Mahe, M.F.; Ribadeau-Dumas, B.; Mercier, J.C.; Grosclaude, F. Two of the Three Genetic Variants of Goat α 1-casein Which are Synthesized at a Reduced Level Have an Internal Deletion Possibly Due to Altered RNA Splicing. *Eur. J. Biochem.* **1990**, *193*, 237–241. [\[CrossRef\]](#)
104. 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. [\[CrossRef\]](#)
105. Caira, S.; Pinto, G.; Balteanu, V.A.; Chianese, L.; Addeo, F. A Signature Protein-Based Method to Distinguish Mediterranean Water Buffalo and Foreign Breed Milk. *Food Chem.* **2013**, *141*, 597–603. [\[CrossRef\]](#)
106. Stepic, S.; Kuveljic, J.; Zivotic, I.; Perisic, P.; Bogdanovic, V.; Ignjatovic, A.; Zivkovic, M. CSN1S1 and CSN3 Genetic Variants Affect Milk Quality Traits in the Buffalo Population in Serbia. *J. Anim. Sci. Technol.* **2024**, *74*, 219–229. [\[CrossRef\]](#)
107. Rullo, R.; Caira, S.; Nicolae, I.; Marino, F.; Addeo, F.; Scaloni, A. A Genotyping Method for Detecting Foreign Buffalo Material in Mozzarella di Bufala Campana Cheese Using Allele-Specific- and Single-Tube Heminested-Polymerase Chain Reaction. *Foods* **2023**, *12*, 2399. [\[CrossRef\]](#)
108. Kappeler, S.; Farah, Z.; Puhani, Z. Sequence Analysis of Camelus Dromedarius Milk Caseins. *J. Dairy Res.* **1998**, *65*, 209–222. [\[CrossRef\]](#)
109. Giambra, I.J.; Erhardt, G. Molecular Genetic Characterization of Ovine CSN1S2 Variants C and D Reveal Further Important Variability within CSN1S2. *Anim. Genet.* **2012**, *43*, 642–645. [\[CrossRef\]](#)
110. Cosenza, G.; Pauculio, A.; Feligini, M.; Coletta, A.; Colimoro, L.; Di Berardino, D.; Ramunno, L. A Point Mutation in the Splice Donor Site of Intron 7 in the α 2-Casein Encoding Gene of the Mediterranean River Buffalo Results in an Allele-Specific Exon Skipping. *Anim. Genet.* **2009**, *40*, 791. [\[CrossRef\]](#)
111. Boissard, M.; Hue, D.; Bouniol, C.; Mercier, J.-C.; Gaye, P. Multiple mRNA Species Code for Two Non-allelic Forms of Ovine α 2-casein. *Eur. J. Biochem.* **1991**, *201*, 633–641. [\[CrossRef\]](#)
112. Lagonigro, R.; Pietrola, E.; D’Andrea, M.; Veltri, C.; Pilla, F. Molecular Genetic Characterization of the Goat α 2-Casein E Allele. *Anim. Genet.* **2001**, *32*, 391–393. [\[CrossRef\]](#)
113. 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. [\[CrossRef\]](#)
114. Erhardt, G.; Jäger, S.; Budelli, E.; Caroli, A. Genetic Polymorphism of Goat α 2-Casein (CSN1S2) and Evidence for a Further Allele. *Milchwissenschaft* **2002**, *57*, 137–140. [\[CrossRef\]](#)

115. 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. [\[CrossRef\]](#)
116. Deepika, S.; Gautam, D.; Meena, S.; Ali, M.; Meena, A.S.; Vats, A.; Verma, M.; Rout, P.K.; De, S. Heterogeneity and Diversified Distribution of α S2 Casein Variants in Indian Goats. *Small Rumin. Res.* **2021**, *204*, 106501. [\[CrossRef\]](#)
117. Ibeagha-Awemu, E.M.; Prinzenberg, E.M.; Jann, O.C.; Lühken, G.; Ibeagha, A.E.; Zhao, X.; Erhardt, G. Molecular Characterization of Bovine CSN1S2*B and Extensive Distribution of Zebu-Specific Milk Protein Alleles in European Cattle. *J. Dairy Sci.* **2007**, *90*, 3522–3529. [\[CrossRef\]](#)
118. Cieslak, J.; Pawlak, P.; Wodas, L.; Borowska, A.; Stachowiak, A.; Puppel, K.; Kuczynska, B.; Luczak, M.; Marczak, L.; Mackowski, M. Characterization of Equine CSN1S2 Variants Considering Genetics, Transcriptomics, and Proteomics. *J. Dairy Sci.* **2016**, *99*, 1277–1285. [\[CrossRef\]](#)
119. Brinkmann, J.; Koudelka, T.; Keppler, J.K.; Tholey, A.; Schwarz, K.; Thaller, G.; Tetens, J. Characterization of an Equine α S2-Casein Variant Due to a 1.3 Kb Deletion Spanning Two Coding Exons. *PLoS ONE* **2015**, *10*, e0139700. [\[CrossRef\]](#)
120. Kusza, S.; Veress, G.; Kukovics, S.; Jávör, A.; Sanchez, A.; Angiolillo, A.; Bosze, Z. Genetic Polymorphism of α S1- and α S2-Caseins in Hungarian Milking Goats. *Small Rumin. Res.* **2007**, *68*, 329–332. [\[CrossRef\]](#)
121. Gigli, I.; Maizon, D.O.; Riggio, V.; Sardina, M.T.; Portolano, B. Short Communication: Casein Haplotype Variability in Sicilian Dairy Goat Breeds. *J. Dairy Sci.* **2008**, *91*, 3687–3692. [\[CrossRef\]](#) [\[PubMed\]](#)
122. Naowanat, N.; Suwattana, D.; Suphap, B. Casein Gene Polymorphism of Goat in Thailand. *Thai J. Vet. Med.* **2024**, *54*, 211–219. [\[CrossRef\]](#)
123. Marletta, D.; Bordonaro, S.; Guastella, A.M.; D’Urso, G. Genetic Polymorphism at CSN1S2 Locus in Two Endangered Sicilian Goat Breeds. *J. Anim. Breed. Genet.* **2004**, *121*, 52–56. [\[CrossRef\]](#)
124. Caroli, A.; Chiatti, F.; Chessa, S.; Rignanes, D.; Bolla, P.; Pagnacco, G. Focusing on the Goat Casein Complex. *J. Dairy Sci.* **2006**, *89*, 3178–3187. [\[CrossRef\]](#)
125. Vacca, G.M.; Ouled Ahmed Ben Ali, H.; Pazzola, M.; Sanna, M.; Dettori, M.L.; Carcangiu, V. An Investigation on Allele Frequency at the CSN1S2 Locus and its Relationship with Milk Parameters in the Sarda Goat. *J. Anim. Feed. Sci.* **2009**, *18*, 628–637. [\[CrossRef\]](#)
126. Palmeri, M.; Mastrangelo, S.; Sardina, M.T.; Portolano, B. Genetic Variability at α S2- Casein Gene in Girgentana Dairy Goat Breed. *Ital. J. Anim. Sci.* **2014**, *13*, 116–118. [\[CrossRef\]](#)
127. Dincel, D.; Ardicli, S.; Samli, H.; Vatanserver, B.; Balci, F. 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. Rev.* **2021**, *27*, 91–96.
128. Bozkaya, F.; Mundan, D.; Karabulut, O.; Yerturk, M.; Gurler, S.; Aral, F. An Investigation on the Distribution of 0 and D Alleles of the CSN1S2 Gene in Goat Populations Raised in Southeastern Region of Turkey. *Small Rumin. Res.* **2008**, *78*, 193–196. [\[CrossRef\]](#)
129. Sztankóová, Z.; Mátllová, V.; Kysel’ová, J.; Jandurová, O.M.; Říha, J.; Senese, C. Polymorphism of Casein Cluster Genes in Czech Local Goat Breeds. *J. Dairy Sci.* **2009**, *92*, 6197–6201. [\[CrossRef\]](#)
130. Correddu, F.; Serdino, J.; Manca, M.G.; Cosenza, G.; Pauciuolo, A.; Ramunno, L.; Macciotta, N.P.P. Use of Multivariate Factor Analysis to Characterize the Fatty Acid Profile of Buffalo Milk. *J. Food Compos. Anal.* **2017**, *60*, 25–31. [\[CrossRef\]](#)
131. Attanasio, C.; David, A.; Neerman-Arbez, M. Outcome of Donor Splice Site Mutations Accounting for Congenital Afibrinogenemia Reflects Order of Intron Removal in the Fibrinogen Alpha Gene (FGA). *Blood* **2003**, *101*, 1851–1856. [\[CrossRef\]](#)
132. Cosenza, G.; Illario, R.; Gallo, D.; Riccio, R.; Di Berardino, D.; Ramunno, L. A Preliminary Analysis of mRNA Transcribed from the N Allele at the CSN1S1 Locus of the Goat. *Ital. J. Anim. Sci.* **2003**, *2*, 10–12. [\[CrossRef\]](#)
133. Chatchatee, P.; Järvinen, K.M.; Bardina, L.; Beyer, K.; Sampson, H.A. Identification of IgE- and IgG-Binding Epitopes on α S1-Casein: Differences Patients with Persistent and Transient Cow’s Milk Allergy. *J. Allergy Clin. Immunol.* **2001**, *107*, 379–383. [\[CrossRef\]](#)
134. Ryskaliyeva, A.; Henry, C.; Miranda, G.; Faye, B.; Konuspayeva, G.; Martin, P. Alternative Splicing Events Expand Molecular Diversity of Camel CSN1S2 Increasing its Ability to Generate Potentially Bioactive Peptides. *Sci. Rep.* **2019**, *9*, 5243. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.