



Contents lists available at ScienceDirect

Meta Gene



# Cloning and comparative analysis of gene structure in promoter site of alpha-s1 casein gene in Naeinian goat and sheep



Mojtaba Najafi\*, Ghodrat Rahimi Mianji, Zarakht Ansari Pirsaraie

Department of Animal Science, Sari Agriculture sciences and Natural Resources University, Iran

## ARTICLE INFO

### Article history:

Received 14 July 2014

Revised 14 July 2014

Accepted 2 November 2014

Available online 16 November 2014

### Keyword:

Polymorphism

Transcriptional factors

Naeinian goat and sheep

## ABSTRACT

The 5' end or alpha-S1 casein promoter has a significant role in milk protein gene expression. The understanding of the translation process of alpha-S1 casein mutants will provide us an opportunity to make the best selection in livestock providing more proteins in milk. Blood samples were taken from three hundred of Naeinian goats and sheep, and DNA extraction was done using modified salting out method. Polymerase chain reactions (PCR) were carried out using a specific primer pairs for amplification a fragment of 1133 bp from part of 5'-UTR and exon 1 of alpha s1 casein gene. The *AluI* and *HinfI* restriction enzyme treatment of all samples provided the same homozygous AA genotype in both species. Subsequently, one sample of each species was selected and cloned, and the final sequences were analyzed by BioEdit, CLC genomic, Mega4 and DNASIS MAX software. Several polymorphisms are recognized between Naeinian goat and sheep that are presented on motif sites. In this research, the interested location, including exon I and a part of 5', was analyzed, and genetic element comparisons were done between Naeinian goat and sheep. The number and location of probable binding sites can have a crucial role as a result of antagonistic and synergistic effects on gene regulation activities.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

## Introduction

Among the calcium-sensitive caseins, the  $\alpha$ s1 fraction is undoubtedly the absolute most extensively investigated in livestock. The goat caseins ( $\alpha$ s1-,  $\beta$ -,  $\alpha$ s2 - and  $\kappa$ -casein) are coded by the single autosomal genes,

\* Corresponding author. Tel.: +98 9132625020.

E-mail addresses: [Mojtaba\\_najafy@yahoo.com](mailto:Mojtaba_najafy@yahoo.com) (M. Najafi), [rahimimianji@yahoo.com](mailto:rahimimianji@yahoo.com) (G. Rahimi Mianji), [Zarakht\\_ansari@yahoo.com](mailto:Zarakht_ansari@yahoo.com) (Z. Ansari Pirsaraie).

CSN1S1, CSN2, CSN1S2 and CSN3, which cluster in a DNA segment of approximately 250 kbp, mapped to chromosome 6 (Rijnkels, 2002). All casein fractions are classified by high variability (Sacchi et al., 2005). So far, at least 17 alleles (A, B1, B2, B3, B4, C, D, E, F, G, H, I, L, M, N, O1 and O2) have already been identified, which are related to various levels of CSN1S1 expression in the milk. The CSN1S1 variants are sorted into 4 groups due to the basis of the milk casein alpha-S1 content: strong alleles (A, B1, B2, B3, B4, C, H, L and M) producing almost 3.5 g/l of casein alpha-S1/each; intermediate alleles (E and I; 1.1 g/l); weak alleles (D, F and G; 0.45 g/l); and null alleles (O1, O2, and N) producing no  $\alpha$ S1-casein (Bevilacqua et al., 2002; Cosenza et al., 2008; Grosclaude, 1999; Leroux et al., 1990; Ramunno et al., 2004, 2005; Rando et al., 2000) (Grosclaude, 1999) (Leroux et al., 1992; Mahé and Grosclaude, 1989; Pérez et al., 1994). A recent study showed that the high frequency of the strong genotypes is associated with the production of high fat and protein content milk and with optimal technological properties (Mastrangelo et al., 2013). CSN1S1 has been characterized by 19 exons ranging in size from 24 (exons 5, 6, 7, 8, 10, 13 and 16) to 382/388 bp (exon 19) expanded over 17.5 kb (Pérez et al., 1994). The understanding of the translational process of alpha-S1 casein gene mutants will provide new opportunities to select the best dairy livestock for the preferred milk protein producing genotype. It has been suggested that a mutation, occurring at position of – 1319 of the promoter region, creates an extra putative activator protein (AP-1) binding motif in the sequence of the F allele, which can be responsible for the different expression of F and N alleles (Ramunno et al., 2005). Therefore, it is possible that using these methods, more quick and accurate selection of animal types independently of age, sex and lactation could be done in order to the different expression levels of the single alleles; thus, animals producing milk with specific chemical–physical and technological characteristics could be chosen. According to Ramunno et al. (2005), comparative analysis of the first 200 bp of the CSN1S1 promoter regions of different species showed a homology between goat and other ruminants (similarities of about 96% with cattle, sheep and yak) stronger than that observed with non-ruminants (similarities of about 88% with rabbit, 80.5% with human and 77% with rat). Common expression features suggest the existence of common motifs and a set of transcription factor binding sites in promoter regions of these genes. Multiply alignment of three to eight milk protein gene promoters showed the existence of these common motifs (Lee et al., 1987) (Laird et al., 1988); however, the question is which of those motifs are specific for the whole group of milk protein gene promoters. To date, eight transcription factors important for milk protein gene expressions and the locations of their binding sites in some of these gene promoters have been described as (1) mammary gland-specific nuclear factor (MGF), as a member of the signal transducers and activators of transcription family (STAT5) (Wakao et al., 1995), which also appears to be identical to the milk protein-binding factor (MPBF) (Burdon et al., 1994); (2) mammary cell-activating factor (MAF), a member of Ets-related proteins (Welte et al., 1994a); (3) pregnancy-specific mammary nuclear factor (PMF) (Lee and Oka, 1992); (4) CCAAT: enhancer binding protein (C: EBP) (Raught et al., 1995); (5) CTF:NF1 (Li and Rosen, 1994); (6) single-stranded DNA-binding transcriptional regulator (STR) (Altiok and Groner, 1994); (7) yinyang (YY1) (Meier and Groner, 1994); and (8) glucocorticoid receptor (GR) (Welte et al., 1994b). Transcription factors are recognized sequences mostly with 6–16 bp long (Wingender, 1993). In fact, progesterone receptor (PR), activating protein-1 (AP-1), CCAAT/enhancer-binding protein (C/EBP), signal transducer and activator of transcription 5 (STAT5, originally identified as milk protein binding factors, MPBF) and pregnancy-specific mammary nuclear factor (PMF) are involved in transcriptional activation, whereas yin–yang (YY1) is involved in gene repression activity. Thus, the occurrence of various consensus in gene promoters can indicate which transcription factors may be involved in a gene regulation activity. Previous computer analysis of milk protein gene promoters has shown their more complex structure (Malewski and Zwierzchowski, 1995). Promoter analysis has been shown that gene expression could be affected by transcriptional factors and the study about transcriptional factor binding element improves our knowledge in gene expression. This study was for characterizing of gene structure in upstream site of 5'-flanking region and exon 1 of  $\alpha$ s1-casein gene in Naeanian goat and sheep.

## Materials and methods

### *Animals and DNA extraction*

Blood samples were randomly collected from 300 Naeanian sheep ( $n = 150$ ) and goat ( $n = 150$ ). Genomic DNA was extracted from blood samples using modified salting out method. Quality and quantity of extracting DNA were measured by agarose gel electrophoresis.

### *Amplification of alpha-s1 casein gene*

The specific primers were designed with Vector NTI software and based on the published sequence of the CSN1S1 gene (GenBank accession numbers: AJ504710), a fragment of 1133 bp was amplified (promoter region and exon 1). The PCR mixture contained 25 pM of each primers (forward primer, 5'-CAAGGAGTTGCAATCAACAAG -3' and reverse primer, 5' -GGCCTTGAATATTCTACCAGA -3'), 200 mM of each dNTP and 2U of XT-Taq DNA polymerase in a final volume of 25 ml. The PCR was performed as follows: denaturation 94 °C for 60 s, annealing 65 °C for 60 s and extension 72 °C for 2 min in 35 cycles. The amplification products were analyzed by electrophoresis on a 1.5% agarose gel, using ethidium bromide staining.

### *Genotyping of CSN1S1 locus by PCR-RFLP*

For genotyping and screening of different alleles of alpha s-1 casein in goat and sheep, the PCR-RFLP procedure developed using *AluI* and *HinfI* restriction enzymes. Then, products were analyzed by electrophoresis on a 2% agarose gel.

### *Cloning and sequencing*

Amplimers representative of the PCR-RFLP patterns of each species were selected to be cloned into a pGEMT-Easy vector (Promega, Madison, WI; Cat. No: A1380) and transformed into competent *Escherichia coli* cells, following the protocol recommended by the manufacturer. Amplimers from these clones and the corresponding genomic DNA were selected for subsequent DNA sequencing. Plasmids from the selected clones were extracted using a High Pure Plasmid Isolation Kit (Roche) and were then sequenced in both directions using universal forward and reverse primers at the Bioneer (Korea) and Millegen (France) companies.

### *DNA sequence analysis*

Identical sequences obtained from at least two clones from sheep and goat, and independent PCR amplification from the same sheep, were subjected to further sequence analysis. Sequence analysis was performed using MEGA 4.0 BioEdit DNASIS MAX and CLC GENOMICS software. The BLAST algorithm was used to search the NCBI GenBank databases (<http://www.ncbi.nlm.nih.gov/>) for homologous sequences. In order for alignment and construction of phylogeny tree, the BioEdit and Mega 4.0 was used, respectively. Also, for motif search and the site for restriction enzyme, DNASIS MAX and CLC genomics software was applied.

## **Results and discussion**

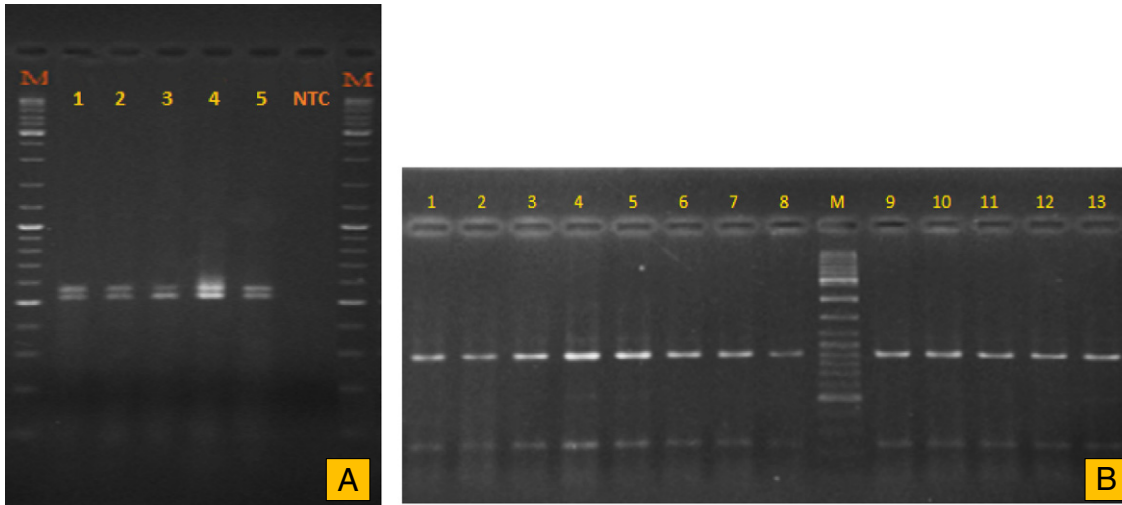
### *Amplification and genotyping*

Polymerase chain reactions were carried out using a specific primer pairs for amplification a fragment of 1133 bp from part of 5'-UTR and exon 1 of alpha s1 casein gene. All extracted DNAs from sheep blood samples yielded a specific single band PCR product without any non-specific band. The *AluI* and *HinfI* restriction enzyme treatment of all samples provided the same homozygous AA genotype (Fig. 1).

Before nucleotide sequencing, one PCR product from each species was purified with high pure PCR product purification kit (Roche; Cat no: 11732668001). Then, the samples were cloned by pGEM T-Easy vector (Promega). The recombinant plasmids were sent to Millegen and Bioneer companies to be sequenced. The gained sequences analyzed and submitted to the NCBI database (accession nos. JN701804.1, JN701803.1 and JN228894.1).

### *Goat CSN1S1 promoter analysis*

In this study, the goat CSN1S1 gene showed two TATA Box and three polyadenylation signals (AATAAA) located, with mention of the first nucleotide of the first exon, at nucleotides -23/-29 and -83/-89 and -586, -457,-420, respectively. Additionally, the 5' end of the gene showed a 16-bp Milk Box motif (CTCCTCAGAATTTCTT, -142/-157) (Schmitt-Ney et al., 1991), a CCAAT/enhancer binding protein (C/EBP



**Fig. 1.** AA genotype samples of *alpha s-1* gene based on PCR-RFLP analyses in Naevian sheep and goat (A (1–5) represents the obtained samples of *AluI*-RE and B (1–13) represents the obtained samples of *HinfI*-RE marker; M represents molecular weight markers; NTC represents non-template control).

						YY1	YY1			
Goat	C	AAGGAGTTGC	AATCAACAAG	TTTTTAAAGC	CCTCACTTGT	ATAGATATTTT	TATTTAGCA	-1031		
Sheep	.	.....	.....	.....	.....	.C.....	.....			
		YY1		YY1						
Goat	C	ATAATATTTT	GTACAATGCC	ATTAATATAT	TGTACAATGT	ACGATGCCAG	TTAACTCTA	-971		
Sheep	.	.....	.....	.....	.....	-----	...T....			
Goat	G	GAGTACAATT	AAGAATTGGA	GGGATAGGAA	TTTTTTTCTT	TTACTTGTTT	ACTTTAAAA	-911		
Sheep	.	.....C.....	...G.....	.....	.....	.....	.....			
				YY1	C/EBP	YY1	AP-1			
goat	G	ATGAAAAATC	AGAGTTATGG	TTTATTTTTT	GCAATATTTA	AAAAATTGTA	TTCTTGAAT	-851		
sheep	.	.....	.....	.....	.....	.....	.....			
goat	A	ATTATTAATT	TTAATTAAT	AATCTGTAAT	GAGAATTCTC	CTACCAATGC	AGGAGACCT	-791		
sheep	.	.....	.....	.....	.....	.....	.....			
				YY1/RC	MAE	PEA3	YY1			
goat	G	AGTTTGATCC	CTGGATAGGG	AAAATA	CCCT	GCAAGAAGGAA	ATGGCAACCC	ACTCCAATA	-731	
Sheep	.	.....	.....	.....	.....	.....	.....			
		YY1								
goat	T	TATATTTGG	GAAATCCCAT	GGACAGAGGA	GACTGGCAGG	CTAAAGTCCA	TGGGGTCAC	-671		
Sheep	.	.....	.....	.....	.....	.....	.....			
					YY1					
goat	A	AAGAACTGGA	CACAACCTAG	TGACTAAACA	ACAACAATTT	ACACCGGAAT	GAATGAAC	-611		
Sheep	.	.....	.....	C.....	.....	.....	...T.....			
			CAAT box	YY1/RC						
goat	A	GTCACCCGAA	CTAGTACACC	CAAAATAAAC	AAAAATTAGC	TTGGTGGTAT	AATTAAAA	-551		
Sheep	.	...T...A...	.....	.....	.....	.....	.....			
				YY1						
goat	G	CCACCAAAGT	TTATACGATA	ATTATATTTT	CTTTTTGCAA	GAAAAAGATT	AGACCACAT	-491		
Sheep	.	.....	.....	.....	.....	.....	.....			
goat	A	TACTGTAAGT	TATTTACAA	GGTAAATAAT	TAAAATAAAT	ATTATGGATT	AACTCAGTT	-431		
Sheep	.	..A.....	.....	.....	.....	.A.....	...G....			
goat	T	TAAGAGGTGA	AATAAATAAT	GAATCTTCT	CATGGTCTTG	TACGTTAATA	AAAATTGAA	-371		
Sheep	.	.....	.....	.....	.....	.....	...G....			
					A5					
goat	A	ATTTTTGAAG	ACCCTGTTTT	GTCCAAGGA	TTTCTTTTAC	AGGTATTGAA	TTTTTCAA	-311		
Sheep	.	.....	.....	.....	.....	.....	.....			
goat	A	GTTACAAAGG	AAATTTTATT	GATATAAATG	CATGTTCTCA	TAATAACCAT	AAATCTAGG	-251		
Sheep	.	.....	.....	A.....	.....	.T.....	.....			
							YY1			
goat	G	TTTTAATGGT	T---TTTTT	TTTTTTTTGG	ATGTTAATTT	AGAACAATGC	CATTCATT	-191		
Sheep	.	...GT...G	.TTTG.....	.....	.....	.....	.....			
			AP-1		Yu-lee 2	Milk box	MGF/STAT5			
goat	T	CCTATATAAT	GAATCACTTC	TITGTTGTA	ACTCTCCTCA	GAATTTCTTG	GGAGAGGAA	-131		
Sheep	.	.....	.....	.....	.....	.....	.....			
			2GRE-1	A4	MPEF/STAT5	TATA box	Yu-lee 4			
goat	C	TGAACAGAAC	ATCGATTTC	TACGTGAGAG	AATCTTAGA	ATTTAAATAA	ACCTATTGG	-71		
Sheep	.	.....	.....	G.T.....	.....	.....	...C....			
		Yu-lee 5	A3	SV40 core enhancer	OCT_1	Oka box	A2	TATA box	Yu-lee 6	A1
goat	T	TAAACTGAAA	CCACAAAAT	AGCATTTC	TAATCAGTAG	GTTAAATAG	CTCGGAAGC	-11		
Sheep	.	.....	.....	.....	.....	.....	...A.....			
			+1	PMF						
goat	A	AAAGTCTGCC	ATCACCTGA	TCATCAACCC	AGCTTGCTGC	TTCTCCCTG	TCTTGGGTT	49		
Sheep	.	.....	.....	.....	.....	.....	.....			

**Fig. 2.** Homology between the nucleotide (nt) sequences of the 5'- flanking region and exon 1 of goat (upper line) and sheep CSN1S1 gene. Numbering is relative to the first nucleotide of the first exon (+ 1) and dashes represent nt identical to those in upper lines. RC: reverse complement. Congruent and putative factors are double-underlined and in shaded bold letters, respectively. Motifs typical for milk protein gene promoters (designed A1–A6) are in italic letters. Abbreviations: YY1, yin and yang common factor 1; PMF, pregnancy-specific mammary nuclear factor; MPEF/STAT5, milk protein binding factor; MAE, mammary cell-activating factor; C/EBP, CCAAT/enhancer-binding protein; NF Oct-1, nuclear factor octamer-1; PR, progesterone receptor; AP-1, activator protein.

–876/–884) sites (Raught et al., 1995), a mammary cell-activating factor (MAF, –750/–758) (Welte et al., 1994b) and a putative binding site (GAATCTTAGAATT) for STAT5 –101/–88), which mediates prolactin signal transduction in lactating mammary gland (Wakao et al., 1995). Other DNA *cis*-acting elements, such as the Simian Virus 40 (SV40)-type enhancer (Weiher et al., 1983), a nuclear factor octamer-1 (NF Oct-1) site (Bohmann et al., 1987) and a YY1 (yin and yang factor 1) common factor (Seto et al., 1991), are spread over the large sequence AAACCACARAATTAGCATNTT (–43/–64), which is conserved in most calcium sensitive casein genes and represents one of the motifs typical for milk protein gene promoters (Malewski, 1998). As presently reported for the bovine CSN1S1 gene (Schild and Geldermann, 1996), a binding site for the pregnancy-specific mammary nuclear factor (PMF) (Lee and Oka, 1992) (nucleotides +8/+16) and two for the activator protein (AP-1) (nucleotides –850/–856 and –175/–181) (Lee et al., 1987) have already been observed. At least another 16 probable YY1 and two NF Oct-1 binding sites could be observed (Fig. 2). The ovine  $\alpha$ S1-CSNGP includes a sequence TTAAATA at –29, which shows homology with the TATA box of bovine, buffalo, camel, rat  $\alpha$ S1-casein genes and  $\beta$ -casein gene of rat. Another TATA sequence is found at 60 bp upstream of the first TATA box site. At the 2nd TATA box, the sequence TTAAATA exists, but the canonical 5-G and 3-AG nucleotides surrounding many eukaryotic TATA sequences (Breathnach and Chambon, 1981) are missing, while the 5-G and 3-AG sequences are present in the first TATA box, which was similar to the result of Bhure and Sharma (2007). The other motifs are spread through the ovine  $\alpha$ S1-casein 5'-flanking region. However, the majority of the milk gene promoter specific motifs is clustered between transcription initiation site and 2155 with the exception of MAF and an upstream MGF/STAT5 site. The most popular motifs occurring in various milk protein gene promoters suggest that their expression is regulated by specific transcription factor(s) and/or by a pattern of transcription factors. However, the computer analysis can predict and localize the binding motifs of transcription factors in promoters creating the experimentation simpler, but the expression studies are needed to confirm the predictions. STAT5 is the most important one of milk protein transcription factors. This factor could be activated by prolactin activation and growth hormone by Jak/STAT signaling pathway or by Src.kinases (EGF receptor) (Gallego et al., 2001). In this study, it has been revealed that nucleus protein bindings for two MGF/STAT5 binding sites, that located in the promoter region of  $\alpha$ S1 casein gene (–90, –98 and –138, –146), are different in one single nucleotide polymorphism (G/A), and these results were completely in agreement with Szymanowska et al. (2004). MGF/STAT5 was demonstrated to be the main TF-mediating prolactin activity on the milk protein gene expressions. The STAT5 DNA-binding capacity was revealed to improve lactation in mammary glands of rodents (Liu et al., 1996) and rabbit (Malewski and Zwierzchowski, 2002). Furthermore, the computer analysis with the TESS program revealed the clear presence of putative binding sites for CP2 (Szymanowska et al., 2004) and G6 factors (Goding et al., 1987). Nucleotide T/C substitution at –733 position in the promoter site of  $\alpha$ S1-casein gene reduced its sequence similarity to CP2 and improved that to SP1 and G6 binding sites. In our study, none of these binding sites was observed.

The existence of six common superfamily-specific motifs of milk protein gene promoters (designed as an A1–A6), known by the way of a complete homology with the consensus sequences described by Malewski (1998), was confirmed in the present study. Motifs A5 and A6 are observed in a distal region apart from A1 to A4 motifs, which are found nearby the start of transcription. The region spanning A1 to A4 motifs includes binding sites for at least six transcription factors that could affect expression and regulation of goat  $\alpha$ S1-casein in the lactating mammary gland. As described earlier, so far, eight transcription factors that are very important in the milk protein gene expression and the locations of their binding sites in some of these gene promoters have already been identified. The results from comparative analysis of Naevian sheep and goats showed the existence of the several SNP between them, which were due to differences in present or non-present in TFBS, and also, in number of transcription factor binding sites. For example, AP-1 C-Myb, CF-1, GATA-1, MGF, Myb, V-MCS, C/EBP, CAP-site, gamma-IRE, NF-IL-6 and HIS4-US motif sites could be mentioned that may be effective on milk protein gene expression. Calvo et al. (2013) reported two polymorphisms in the 5' flanking region that were located within possible trans-acting factor binding sites, modifying the putative CdxA and GATA-1 consensus sites (Calvo et al., 2013). Also, the SNP (JN560175: g. 1123C > A) that modifies a putative CdxA consensus site showed a significant effect on the expression of the CSN1S1 gene ( $P = 0.043$ ), but no effect was found in protein content. This study was not observed, named SNP (this SNP position is shown by highlighting nucleotide at position –41). Promoter regions of milk protein genes are commonly used to produce pharmaceutically and medically important proteins in the mammary gland of transgenic animals and also could be utilized for the construction of an inducible eukaryotic expression vector. In this field, a study on cloning and characterization of ovine  $\alpha$ S1-casein gene promoter DNA sequence analysis showed

the presence of all the core promoter elements such as milk specific transcription factor sequence (MBPF, MGF/STAT5, Yu Lee2, Yu Lee4, Yu Lee5, Yu Lee6 and Oka Box C), prolactin, progesterone responsive elements and one glucocorticoid responsive element and another promoter element that was in consistence with our results at IRE (Bhure and Sharma, 2007).

## Conclusion

In this study, we analyzed 5-flanking region and exon 1 of  $\alpha$ s1 casein gene in Naeinian goats and sheep, which is often used for identification and evaluation of the role and importance of transcription factors in gene expression regulation and transcription effects from polymorphisms that located in this region. However, these sequences usually present in everywhere, but sequence intervals and assortment are different between various species of promoters. The existence of common motifs in various milk protein gene promoters showed that gene expression is regulated by transcription factors. Although, computer analysis could be predicted and localized the binding motifs of transcription factors in promoters making the experimentation simpler, but the expression studies are required to confirm the predictions. Characterization of regulated transcription sequences in gene promoter regions is a very important factor for milk protein production in a variety of species. The aim of this study was to clone, sequence and identify regulating elements in  $\alpha$ s1 casein gene and also to detect promoter nucleus DNA elements (*cis*-acting), TATA signal, GC box, CAAT box and other elements in goats and sheep. The understanding of the translation process of alpha-S1 casein polymorphisms provide us an opportunity to do the most effective selection in livestock for the preferred milk protein producing genotype. The results of the present study provided important evidences for the merits of the promoter to construct a eukaryotic expression vector and a series of information for transgenic studies.

## References

- Altiok, S., Groner, B., 1994. beta-Casein mRNA sequesters a single-stranded nucleic acid-binding protein which negatively regulates the beta-casein gene promoter. *Molecular and Cellular Biology* 14, 6004–6012.
- Bevilacqua, C., Ferranti, P., Garro, G., Veltri, C., Lagonigro, R., Leroux, C., Pietrolà, E., Addeo, F., Pilla, F., Chianese, L., 2002. Interallelic recombination is probably responsible for the occurrence of a new  $\alpha$ s1-casein variant found in the goat species. *European Journal of Biochemistry* 269, 1293–1303.
- Bhure, S., Sharma, B., 2007. Cloning and characterization of ovine  $\alpha$ S1-casein gene promoter: a transfection study in rat mammary gland cell line: full length research paper. *Mitochondrial DNA* 18, 39–46.
- Bohmann, D., Keller, W., Dale, T., Schöler, H.R., Tebb, G., Mattaj, I.W., 1987. A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. *Nature* 325, 268–272.
- Breathnach, R., Chambon, P., 1981. Organization and expression of eucaryotic split genes coding for proteins. *Annual Review of Biochemistry* 50, 349–383.
- Burdon, T.G., Demmer, J., Clark, A.J., Watson, C.J., 1994. The mammary factor MPBF is a prolactin-induced transcriptional regulator which binds to STAT factor recognition sites. *FEBS Letters* 350, 177–182.
- Calvo, J., Dervishi, E., Sarto, P., González-Calvo, L., Berzal-Herranz, B., Molino, F., Serrano, M., Joy, M., 2013. Structural and functional characterisation of the  $\alpha$ -S1-casein (<i>CSN1S1</i>) gene and association studies with milk traits in Assaf sheep breed. *Livestock Science* 157, 1–8.
- Cosenza, G., Pauciullo, A., Gallo, D., Colimoro, L., D'Avino, A., Mancusi, A., Ramunno, L., 2008. Genotyping at the <i>CSN1S1</i> locus by PCR-RFLP and AS-PCR in a Neapolitan goat population. *Small Ruminant Research* 74, 84–90.
- Gallego, M.I., Binart, N., Robinson, G.W., Okagaki, R., Coschigano, K.T., Perry, J., Kopchick, J.J., Oka, T., Kelly, P.A., Hennighausen, L., 2001. Prolactin, growth hormone, and epidermal growth factor activate Stat5 in different compartments of mammary tissue and exert different and overlapping developmental effects. *Developmental Biology* 229, 163–175.
- Goding, C., Temperley, S., Fisher, F., 1987. Multiple transcription (actors interact with the adenovirus-2 EII-late promoter: evidence for a novel CCAAT recognition factor. *Nucleic Acids Research* 15, 7761–7780.
- Grosclaude, F., 1999. Genetic polymorphism of milk proteins.
- Laird, J., Jack, L., Hall, L., Boulton, A., Parker, D., Craig, R., 1988. Structure and expression of the guinea-pig alpha-lactalbumin gene. *Biochemical Journal* 254, 85–94.
- Lee, C.S., Oka, T., 1992. A pregnancy-specific mammary nuclear factor involved in the repression of the mouse beta-casein gene transcription by progesterone. *Journal of Biological Chemistry* 267, 5797–5801.
- Lee, W., Mitchell, P., Tjian, R., 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* 49, 741–752.
- Leroux, C., Martin, P., MAHÉ, M.F., Levéziel, H., MERCIER, J.C., 1990. Restriction fragment length polymorphism identification of goat  $\alpha$ s1-casein alleles: a potential tool in selection of individuals carrying alleles associated with a high level protein synthesis. *Animal Genetics* 21, 341–351.
- Leroux, C., Mazure, N., Martin, P., 1992. Mutations away from splice site recognition sequences might cis-modulate alternative splicing of goat alpha s1-casein transcripts. Structural organization of the relevant gene. *Journal of Biological Chemistry* 267, 6147–6157.
- Li, S., Rosen, J.M., 1994. Glucocorticoid regulation of rat whey acidic protein gene expression involves hormone-induced alterations of chromatin structure in the distal promoter region. *Molecular endocrinology* (Baltimore, Md.) 8, 1328–1335.
- Liu, X., Robinson, G.W., Hennighausen, L., 1996. Activation of Stat5a and Stat5b by tyrosine phosphorylation is tightly linked to mammary gland differentiation. *Molecular Endocrinology* (Baltimore, Md.) 10, 1496–1506.

- Mahé, M., Grosclaude, F., 1989. Synthesis rate at the caprine  $\alpha$ S1-casein locus.
- Malewski, T., 1998. Computer analysis of distribution of putative cis- and trans-regulatory elements in milk protein gene promoters. *Biosystems* 45, 29–44.
- Malewski, T., Zwierzchowski, L., 1995. Computer-aided analysis of potential transcription-factor binding sites in the rabbit  $\beta$ -casein gene promoter. *Biosystems* 36, 109–119.
- Malewski, T., Zwierzchowski, L., 2002. Genes expressed in cattle mammary gland—computational analysis of 5'-upstream sequences in search for factors conferring a tissue- and stage-specific transcription. *Animal Science Papers and Reports* 20, 5–20.
- Mastrangelo, S., Sardina, M., Tolone, M., Portolano, B., 2013. Genetic polymorphism at the CSN1S1 gene in Girgentana dairy goat breed. *Animal Production Science* 53, 403–406.
- Meier, V.S., Groner, B., 1994. The nuclear factor YY1 participates in repression of the beta-casein gene promoter in mammary epithelial cells and is counteracted by mammary gland factor during lactogenic hormone induction. *Molecular and Cellular Biology* 14, 128–137.
- Pérez, M.J., Leroux, C., Bonastre, A.S., Martin, P., 1994. Occurrence of a LINE sequence in the 3' UTR of the goat  $\alpha$ S1-casein E-encoding allele associated with reduced protein synthesis level. *Gene* 147, 179–187.
- Ramunno, L., Cosenza, G., Rando, A., Illario, R., Gallo, D., Berardino, D.D., Masina, P., 2004. The goat  $\alpha$ S1-casein gene: gene structure and promoter analysis. *Gene* 334, 105–111.
- Ramunno, L., Cosenza, G., Rando, A., Pauciuolo, A., Illario, R., Gallo, D., Di Berardino, D., Masina, P., 2005. Comparative analysis of gene sequence of goat CSN1S1 F and N alleles and characterization of CSN1S1 transcript variants in mammary gland. *Gene* 345, 289–299.
- Rando, A., Ramunno, L., Masina, P., 2000. Mutations in casein genes. *Zootecnica e Nutrizione Animale* 26, 105–114.
- Raught, B., Liao, W.-L., Rosen, J.M., 1995. Developmentally and hormonally regulated CCAAT/enhancer-binding protein isoforms influence beta-casein gene expression. *Molecular endocrinology (Baltimore, Md.)* 9, 1223–1232.
- Rijnkels, M., 2002. Multispecies comparison of the casein gene loci and evolution of casein gene family. *Journal of Mammary Gland Biology and Neoplasia* 7, 327–345.
- Sacchi, P., Chessa, S., Budelli, E., Bolla, P., Ceriotti, G., Soglia, D., Rasero, R., Cauvin, E., Caroli, A., 2005. Casein haplotype structure in five Italian goat breeds. *Journal of Dairy Science* 88, 1561–1568.
- Schild, T., Geldermann, H., 1996. Variants within the 5'-flanking regions of bovine milk-protein-encoding genes. III. Genes encoding the Ca-sensitive caseins  $\alpha$ S1,  $\alpha$ S2 and  $\beta$ . *Theoretical and Applied Genetics* 93, 887–893.
- Schmitt-Ney, M., Doppler, W., Ball, R.K., Groner, B., 1991. Beta-casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor. *Molecular and Cellular Biology* 11, 3745–3755.
- Seto, E., Shi, Y., Shenk, T., 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. *Nature* 241–245.
- Szymanowska, M., Malewski, T., Zwierzchowski, L., 2004. Transcription factor binding to variable nucleotide sequences in 5'-flanking regions of bovine casein genes. *International Dairy Journal* 14, 103–115.
- Wakao, H., Gouilleux, F., Groner, B., 1995. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *The EMBO Journal* 14, 854.
- Weiber, H., König, M., Gruss, P., 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* 219, 626–631.
- Welte, T., Garimorth, K., Philipp, S., Jennewein, P., Huck, C., Cato, A.C., Doppler, W., 1994a. Involvement of Ets-related proteins in hormone-independent mammary cell-specific gene expression. *European Journal of Biochemistry* 223, 997–1006.
- Welte, T., Garimorth, K., Philipp, S., Doppler, W., 1994b. Prolactin-dependent activation of a tyrosine phosphorylated DNA binding factor in mouse mammary epithelial cells. *Molecular endocrinology (Baltimore, Md.)* 8, 1091–1102.
- Wingender, E., 1993. Gene regulation in eukaryotes. VCH.