# GENETIC POLYMORPHISMS OF $\alpha_{S1}$-CASEIN (*CSN1S1*) AND β-CASEIN (*CSN2*) GENES IN CARPATHIAN GOAT BREED

**V.A. Balteanu[1*], C. Pascal[2], A. Vlaic[1]**

[1]*University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Romania*
[2]*University of Agricultural Sciences and Veterinary Medicine Iasi, Romania*

**Abstract**

*In goat casein loci a large genetic variation was described with up to 20 alleles found at the $\alpha_{S1}$-casein (CSN1S1) locus and 9 alleles at the β-casein (CSN2) locus. In Carpathian goat breed little is known about the nucleotide variations occurring at the CSN1S1 and CSN2 loci. Therefore the objective of this study was to identify the types of alleles occurring at these loci in a Carpathian goat population by using proteomic and genomic analysis tools. Isoelectric focusing (IEF) analysis of milk samples allowed the identification of some CSN1S1 variants (A, B/E, C and F) based on isoelectric point differences. Furthermore AS-PCR genotyping test allowed the discrimination of E from non-E alleles. The sequencing of CSN1S1 and CSN2 cDNA obtained from some goats from the analysed population revealed that the IEF bands attributed to CSN1S1 B contains two variants (B2 and B4), while the monomorphic IEF CSN2 pattern contains the A and C variants. However the variation of CSN1S1 and CSN2 loci in Carpathian goat breed could be more complex and therefore further studies on a larger number of goats are needed to characterize it.*

**Key words**: goat, milk, casein, polymorphisms, alleles

## INTRODUCTION

In ruminants, casein micelles are composed of three calcium sensitive caseins: $\alpha_{S1}$-casein (CSN1S1), β-casein (CSN2) and $\alpha_{S2}$-casein, which are stabilized by k-casein (CSN3) [1]. In goats, CSN1S1 and CSN2 are encoded by *CSN1S1* and *CSN2* genes, mapping to chromosome 6 [1, 2] and spaced 12 kb apart [3].

At least 9 alleles are currently characterized at the goat *CSN2* locus [4, 5], which are associated either with normal CSN2 content (5g/L/allele) or with absence of CSN2 in milk [6-8]. The most remarkable polymorphism is found at the goat *CSN1S1* locus, with at least 20 alleles currently known [4, 9, 10], which are associated with four different expression levels ranging from 0 to 3.5 g/L/allele [11]. This genetic variation found at the goat *CSN1S1* locus has a significant effect on milk quality, rheological properties or cheese yield and it is well documented in various breeds [12-16].

The correct identification of this genetic variation represents an essential step to conceive breeding strategies aiming to improve milk quality in a particular goat breed. The traditional isoelectric focusing (IEF) technique allows rapid screening and identification of the majority of milk protein variants found in goat, based on isoelectric point (IP) differences. However, silent amino acid substitutions characterizing some of these protein variants, which have no effect on protein IP, makes this approach unreliable. These allelic variations are detectable only by PCR based methods.

Therefore the objective of this study was to identify the types of alleles occurring at the *CSN1S1* and *CSN2* loci in a Carpathian goat population by using proteomic and genomic analysis tools.

## MATERIAL AND METHODS

### Milk sampling and IEF analysis

Milk samples (15 ml/goat) were collected in sterile tubes by hand milking from 50

*Corresponding author: avbalteanu@yahoo.com
*The manuscript was received: 25.03.2015*
*Accepted for publication: 15.04.2015*

Carpathian goats from Transylvanian region and stored at -20 ºC. After defrosting, milk samples were denatured in 8 M urea and 3% dithiothreitol and incubated for 2 h at 4 ºC. Subsequently, denatured milk samples were loaded in 4 % ultrathin (0.5 mm) polyacrylamide gels that contained 8 M urea and a mixture of three ampholytes: pH=2.5-5, pH=4.2-4.9, pH=5.0-7.0 (GE Healthcare, Sweden) [17]. The gels were run in a Multiphor II Electrophoresis System and stained with PhastGel Blue R (GE Healthcare, Sweden). The gels images were captured and analyzed with a Molecular Imager Gel Doc XR System (Bio-Rad, USA).

**Differentiation of *CSN1S1* E allele from non-E alleles by AS-PCR**

The goats exhibiting different CSN1S1 IEF phenotypes (BB, BE or EE) were further analysed for the presence of the 457 bp LINE element specific to E allele [18]. For DNA extraction, blood collected from these goats on $K_3$-EDTA was processed with the Quick-gDNA MiniPrep kit (Zymo Research Corporation, USA). The DNA concentration and purity was determined on a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA).

All PCR reactions were set up at 25 µl final volume that contained 1X Tissue Green PCR Master Mix (Fermentas, Lithuania), 10 pmol of each specific primer reported before [19] and 50 ng of genomic DNA as template in each reaction. The thermal profile was as follows: 94 °C for 3 min followed by 35 cycles of 94 °C for 1 min, 58 °C for 1 min, 72 °C for 1 min and a final extension step of 72 °C for 7 min. The PCR products were analyzed on a 2 % agarose gel containing 1X SybrSafe (Invitrogen, USA) and in 1X TBE buffer (Lonza, Belgium).

**Total RNA isolation, cDNA synthesis and sequencing**

To extract total RNA from milk somatic cells, eight samples exhibiting different IEF patterns for *CSN1S1* were selected. After defrosting milk samples were centrifuged at 2,000 g for 15 min at 4 ºC. The resulting cell pellets were mixed with 1 ml PureZOL reagent (Bio-Rad, USA) and purified according to the manufacturer instructions. The RNA concentration and purity was assessed on a Nanodrop spectrophotometer.

To synthesize first cDNA strand the iScript$^{TM}$cDNA Synthesis Kit (Bio-Rad, USA) was used. To amplify the entire coding region of the goat *CSN1S1* cDNA a set of specifically designed primers reported before [17] was used. For *CSN2* cDNA amplification a set of newly designed primers ChBCZ - F: 5`- CTC CTT CAC TTC TTC TCC TCT - 3`, ChBCZ - R: 5`- TTG CCA TAT TTC CA GTC GC AG - 3) was used. The primers were designed with the Primer3 v.0.4.0 software [20], based on the goat *CSN2* mRNA sequence of A allele (GenBank Acc. No. AJ011018.3).

The amplification mix was prepared in 25 µl final volume that contained 1X Tissue Green PCR Master Mix (Fermentas, Lithuania), 10 pmol of each specific primer pair for the target transcripts (*CSN1S1* or *CSN12*) and 1 µl of each reverse transcription reactions. The PCR cycling and electrophoresis of amplicons were carried out in the same conditions as described in the previous paragraph.

The resulting *CSN1S1* and *CSN12* cDNA amplicons were bi-directionally sequenced by using the same primers and the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, USA). Sequencing products were analyzed by capillary electrophoresis on an Applied Biosystem 3730 device (Applied Biosystems, USA). The chromatograms derived from cDNA sequencing were analyzed with the BioEdit software (http://www.mbio.ncsu.edu/bioedit/bioedit.html) and translated into amino acid sequences using the ExPASy translate tool (http://www.expasy.org/). The resulting sequences were aligned with the CLUSTAL W2 analysis software (www.ebi. ac.uk/tools/msa/clustalw2/).

## RESULTS AND DISCUSSIONS

The IEF analysis of the goat milk samples collected from a Carpathian breed population highlighted as expected a high polymorphism at the *CSN1S1* locus, which for the *CSN2* locus was not evident (Figure 1). The

differentiation of CSN1S1 A, B and C variants in IEF gels was possible due to differences in their electrophoretic behaviour and similar intensity of the protein bands (Figure 1). However the B variant might contain B1, B3, B3 and B4 variants or a combination these, which are not separable in IEF gels due to their similar behaviour.

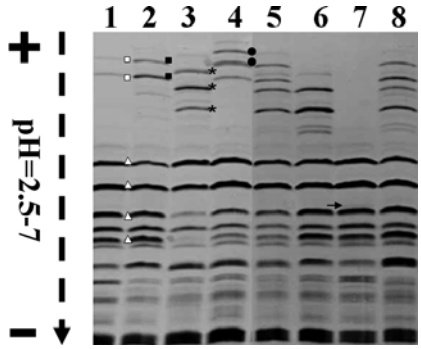


Fig. 1 Isoelectric focusing patterns of some Carpathian goat milk samples separated in a pH range of 2.5-7. CSN1S1 E - white squares; CSN1S1 B - black squares; CSN1S1 A - black stars; CSN1S1 C - black dots; CSN1S1 F - black arrow. CSN1S1 phenotypes: 1 - EE; 2 - BB; 3 -AA; 4 - BC; 5 - AB; 6 - AA; 7 - FF; 8 – AB; CSN2 A/C - white triangles.

CSN1S1 F could be distinguished as a faint band above the 3rd band of CSN2, while CSN1S1 E as faint bands with a similar IP with the B variant (Figure 1). This reduced intensity of the E bands as compared with the B bands is due to the insertion of a LINE element (457 bp in length) within the 19th exon that leads to a reduced protein synthesis rate [18]. However, the correct identification of the E variant based on the intensity of IEF bands is possible only when is present in homozygous condition or in heterozygous condition with other alleles than B. Therefore additional DNA tests were carried out to discriminate the E allele from non-E alleles.

Genotyping results obtained by AS-PCR confirmed is some samples the presence of the LINE element specific to E allele, either in homozygous or heterozygous condition. A single 547 bp DNA band (which includes the 457 LINE element) indicates the presence of E allele, while an additional 90 bp fragment signals the presence of non-E alleles.

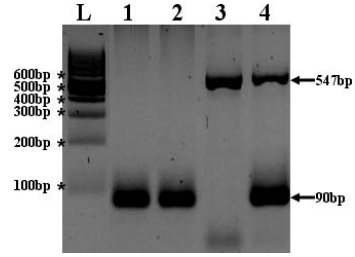


Fig. 2 Discrimination of *CSN1S1* E allele from non-E alleles by AS-PCR amplification of the LINE element located in the 19th exon. L - 100 bp DNA Ladder (Fermentas, Vilnius, Lithuania); Lanes 1 and 2: non-E alleles carriers; Lane 3: E allele carriers in homozygous condition; Lane 4: E allele carriers in heterozygous condition.

As regards the CSN2, four IEF bands with different phosphorylation levels were detected as expected that could be attributed either to A or C variants, which are not separable by IEF due to their similar IP.

Based on the IEF and AS-PCR data the allelic frequency was calculated for the *CSN1S1* and *CSN2* loci (Table 1).

Table 1 Allelic frequency at *CSN1S1* and *CSN2* loci in the analyzed Carpathian goat population

| Locus | No. of goats | Allelic frequency |
|---|---|---|
| *CSN1S1* | 84 | A B* C E F |
| | | 0.38;0.29;0.02;0.15;0.16 |
| *CSN2* | 84 | A+C** |
| | | 1 |

*The value includes the frequencies of B1, B2, B3 or B4 variants that have similar IEF behaviour
**Not distinguished from A variant due to the similar IEF behaviour

In *CSN1S1* locus we found a higher frequency for strong expression alleles (A, B and C) in the two analyzed Carpathian goat populations, as compared with medium and low expression alleles (E and F). In *CSN2* locus the cumulative frequency of A and C alleles was 1 (Table 1).

To further confirm at the DNA level the identity of the observed IEF patterns for CSN1S1 and CSN2 and to identify the

variants that were not separable by IEF but might be present in the analysed goat populations, we comparatively analysed the nucleotide sequences derived from *CSN1S1* and *CSN2* cDNA sequencing of the eight IEF polymorphic samples (shown in Figure 1). Analysis of the cDNA sequences revealed two additional *CSN1S1* alleles B2 and B4, which were previously reported in other breeds [21, 22]. They were assigned as B by IEF analysis due their similar IP. Sequencing data revealed that they are distinguished by two SNP located in exons 12 and 17 and are responsible for two amino acid substitutions *i.e.* p.Arg100Lys and p.Thr195Ala (Table 2).

Table 2 Alleles identified at the *CSN1S1* locus in the analyzed Carpathian goat population

| Locus | Nucleotide variation in mARN / amino acid variation in mature protein | | | | | | | Identified alleles |
|---|---|---|---|---|---|---|---|---|
| *CSN1S1* | exon 3 | exon 4 | exon 9 | exon 10 | exon 12 | exon 17 | exon 19 | |
| | **C**$^{16}$**A**$^{17}$C<br>His$^{8}$ | C**C**$^{8}$C<br>Pro$^{16}$ | T**C**$^{23}$G<br>Ser$^{66}$ | **G**$^{22}$AG<br>Glu$^{77}$ | A**G**$^{14}$A<br>Arg$^{100}$ | **A**$^{139}$CT<br>Thr$^{195}$ | - | B2 |
| | - | - | - | - | A**A**$^{14}$A<br>Lys$^{100}$ | **G**$^{139}$CT<br>Ala$^{195}$ | - | B4 |
| | **A**$^{16}$**T**$^{17}$C<br>Ile$^{8}$ | - | - | - | A**A**$^{14}$A<br>Lys$^{100}$ | **G**$^{139}$CT<br>Ala$^{195}$ | - | C |
| | - | C**T**$^{8}$C<br>Leu$^{16}$ | - | **C**$^{22}$AG<br>Gln$^{77}$ | - | - | - | A |
| | - | - | - | - | A**A**$^{14}$A<br>Lys$^{100}$ | **G**$^{139}$CT<br>Ala$^{195}$ | 457 bp insertion | E |
| | - | - | Deletion of exons 9-11 due to a cytosine deletion at 23$^{rd}$ position of the 9$^{th}$ exon | | - | - | - | F |

Although CSN2 A and C variants were not separable in IEF gels, sequencing data confirmed the presence of both alleles in the analysed Carpathian goat population. These alleles are distinguished by a SNP located in exon 7 that causes at the protein level a p.Ala100Val exchange [23]. This amino acid substitution has no effect on protein IP, which explain the similar IEF behaviour of the A and C variants (Table 2).

Table 2 Alleles identified at the *CSN2* locus in the analyzed Carpathian goat population

| Locus | Nucleotide variation in mARN / amino acid variation in mature protein | Identified allele |
|---|---|---|
| *CSN2* | Exon 7 | |
| | G**C**$^{404}$ A<br>Ala$^{177}$ | A |
| | G**T**$^{404}$ A<br>Val$^{177}$ | C |

To determine the genotyping accuracy determined by IEF analysis of milk samples, we compared these results with those derived from AS-PCR genotyping and cDNA sequencing of the eight mentioned samples (Table 3).

Table 3 Comparison of the genotyping data obtained at the goat *CSN1S1* and *CSN2* loci

| No* | Phenotypes determined by IEF | | Genotypes determined by AS-PCR and sequencing | |
|---|---|---|---|---|
| | CSN1S1 | CSN2 | *CSN1S1* | *CSN2* |
| 1 | EE | AA | EE | AA |
| 2 | BB | AA | $B_4E$ | AA |
| 3 | AA | AA | AA | AC |
| 4 | BC | AA | $B_4C$ | AA |
| 5 | AB | AA | $AB_2$ | AC |
| 6 | AA | AA | AA | AA |
| 7 | FF | AA | FF | AA |
| 8 | AB | AA | $AB_2$ | AA |

* Sample number in IEF gel (Figure 1)

In general, the IEF analysis of milk samples allowed the correct identification of the *CSNS1* variants in majority of the samples. However, for accurate genotyping of E allele an additional PCR test was needed to detect the presence of the LINE element specific to this allele. Furthermore the B2 and

B4 alleles were identifiable by only by sequencing. This was also the case for the *CSN2* A and C alleles (Table 3).

## CONCLUSIONS

We describe in this study the genetic variation occurring at the *CSN1S1* and *CSN2* loci that was assessed in a Carpathian goat population by using a combined approach of protein and DNA analysis. By cDNA sequencing we discriminated two *CSN1S1* alleles (B2 and B4) and two *CSN2* alleles (A and C) that were identifiable neither by AS-PCR nor by IEF analysis of milk samples. We can conclude that for accurate genotyping of the goats at the *CSN1S1* and *CSN2* loci a combination of these methods is needed. However, we analysed in the present study a limited number of samples and the variation of *CSN1S1* and *CSN2* loci in Carpathian goat breed could be more complex. Therefore further studies are needed to characterize it.

## ACKNOWLEDGMENTS

This paper was published under the frame of the European Social Fund, Human Resources Development Operational Program 2007-2013, project no. POSDRU/159/1.5/S/132765.

## REFERENCES

[1] Rijnkels M., Multispecies comparison of the casein gene loci and evolution of casein gene family, J. Mammary Gland Biol., Neoplasia, 2002, 7, p. 327-345.
[2] Threadgill D.W., Womack J.E., Genomic analysis of the major bovine casein genes, Nucleic Acids Res., 1990, 18, p. 6935-6942.
[3] Leroux, C., Martin P., The caprine alphas1 and beta-casein genes are 12-kb apart and convergently transcribed, Anim. Genet., 1996, 27, p. 93.
[4] Marletta D., Criscione A., Bordonaro S., Guastella A.M., D'Urso G., Casein polymorphism in goat's milk, Lait, 2007, 87, p. 491-504.
[5] Chessa S., Rignanese D., Kupper J., Pagnacco G., Erhardt G., Caroli A., The b-casein (CSN2) silent allele C1 is highly spread in goat breeds, J. Dairy Sci., 2008, 91, p. 4433-4436.
[6] Ramunno L., Mariani P., Pappalardo M., Rando A., Capuano M., Di Gregorio P., Cosenza G., A major-effect gene for β-casein content in goat milk, in: Atti XI Congresso Nazionale ASPA, Grado, Italy, 1995, p. 185-186.
[7] Rando A., Pappalardo M., Capuano M., Di Gregorio P., Ramunno L., Two mutations might be responsible for the absence of β-casein in goat milk, Anim. Genet., 1996, p. 27: 31.
[8] Persuy M.A., Printz C., Medrano J.F., Mercier J.C., A single nucleotide deletion resulting in a premature stop codon is associated with marked reduction of transcripts from a goat β-casein null allele, Anim. Genet., 1999, 30, p. 444-451.
[9] Küpper J., Chessa S., Rignanese D., Caroli A., Erhardt G., Divergence at the casein haplotypes in dairy and meat goat breeds, J. Dairy Res., 2010, 77, p. 56-62.
[10] Balteanu V.A., Amills M., Vlaic A, Identification and characterization of a novel defective allele at the goat (*Capra hircus*) alpha s1-casein locus, In: 34th ISAG, China, 2014, p. 132.
[11] Martin P., Ollivier-Bousquet M., Grosclaude F., Genetic polymorphism of caseins: A tool to investigate casein micelle organization, Int. Dairy J., 1999, 9, p. 163-171.
[12] Delacroix-Buchet A., Degas C., Lamberet G., Vassal L., Influence des variants AA et FF de la caséine αs1 caprine sur le rendement fromager et les caractéristiques sensorielles des fromages, Lait, 1996, 76, p. 217-241.
[13] Caravaca F., Carrizosa J., Urrutia B., Baena F., Jordana J., Amills M., Badaoui B., Sanchez A., Angiolillo A., Serradilla J.M., Short communication: Effect of alphaS1-casein (CSN1S1) and kappa-casein (CSN3) genotypes on milk composition in Murciano-Granadina goats, J. Dairy Sci., 2009, 92, p. 2960-2964.
[14] 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 αs1-casein 'null' variant in milk from Norwegian dairy goats-implications for milk composition, micellar size and renneting properties, Dairy Sci. and Tech., 2010, 91, p. 39-45.
[15] Yue X.P., Zhang X.M., Wang W., Ma R.N., Deng C.J., Lan X.Y., Chen H., Li F., Xu X.R., Ma Y., Lei C.Z., The CSN1S1 N and F alleles identified by PCR-SSCP and their associations with milk yield and composition in Chinese dairy goats, Mol. Biol. Rep., 2011, 38, p. 2821-2825.
[16] Balteanu V.A., Serradilla J. M., Vlaic A., Carsai T.C., Amills M., Miresan V., Effect of αS1-casein (CSN1S1) genotypes on milk composition and cheese yield in Carpathian goat breed, XI IGA Conference, Spain, 2012, p. 216.
[17] Balteanu V.A., Carsai T.C., Vlaic A., Identification of an intronic regulatory mutation at the buffalo αS1-casein gene that triggers the skipping of exon 6, Mol. Biol. Rep., 40, 2013, p. 4311-4316.
[18] Jansà-Perez M.J., Leroux C., Bonastre A.S., Martin P., 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, 1994, p. 179-187.
[19] Feligini M., Frati S., Curik V.A., Brambilla A., Parma P., Curik I., Greppi G.F., Enne, G., Caprine αs1-casein polymorphism: characterisation of A, B, E and F variants by means of various biochemical and molecular techniques. Food Technol. Biotech., 2005, 43, p. 123-132
[20] Rozen S., Skaletsky H., Primer3 on the WWW for general users and for biologist programmers, Methods Mol. Biol., 2000, 132, p. 365-386.
[21] Brignon G., Mahé M.F., Grosclaude F., Ribadeau Dumas B., Sequence of caprine αs1-casein and characterization of those of its genetics variants which are synthesized at high level αs1-CnA, B et C, Protein Seq. Data Anal., 1989, 2, p. 181-188.
[22] Grosclaude F., Martin P., Casein polymorphism in the goat. In: IDF Seminar "Milk Protein Polymorphism II", Palmerston North, 1997, p. 241-253.
[23] Neveu C., Mollé D., Moreno F.J., Martin P., Léonil J.: Heterogeneity of caprine betacasein elucidated by RP-HPLC/MS: genetic variants and phosphorilations, J. Protein Chem. 2002, 21, p. 557-567.