Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar
http://dx.doi.org/10.5424/sjar/2013111-3309

Spanish Journal of Agricultural Research 2013 11(1), 80-83
ISSN: 1695-971-X
eISSN: 2171-9292

Short communication. Single nucleotide polymorphisms in the ovine *CSN1S2* gene for α_{S2}-casein

J. A. Padilla[1]*, J. M. Corral[2], A. Quesada[3], M. Izquierdo[2], J. C. Parejo[1] and M. Martínez-Trancón[1]

[1] *Departamento de Producción Animal y Ciencia de los Alimentos. Facultad de Veterinaria. Universidad de Extremadura. Avda. de la Universidad, s/n. 10071 Cáceres. Spain*
[2] *Centro de Investigación La Orden-Valdesequera. Junta de Extremadura. A-5, km 372. 06071 Guadajira (Badajoz). Spain*
[3] *Departamento de Bioquímica, Biología Celular y Genética. Facultad de Veterinaria. Universidad de Extremadura. Avda. de la Universidad, s/n. 10071 Cáceres. Spain*

Abstract

This work describes eighteen SNPs (9 transitions and 9 transversions) found in the sequence spanning from exon 7 to 11 of the CSN1S2 ovine gene, thirteen located at the introns 7 (5 SNPs), 8 (2 SNPs), 9 (1 SNP), and 10 (5 SNPs), and five at the exons 10 (4 SNPs) and 11 (1 SNP). The allele frequencies of the SNPs have been obtained from seventeen animals of Merino breed. Among intronic polymorphisms, the two adjacent SNPs g.[435*C* > *A*; 436*T* > *G*] are located at the pyrimidine rich box near the 3' splice acceptor site of intron 10, which could potentially affect mRNA processing. Analysis of the deduced amino acid sequences from exons 10 and 11 showed that polymorphisms *ex10*-g.153*G* > *T* and *ex10*-g. 155*C* > *T*, which affect the same codon, and *ex11*-g. *504A* > *G,* are non-synonymous substitutions producing the amino acid changes Asp[75] to Tyr[75] and Ile[105] to Val[105], respectively, giving rise to the A and B alleles previously found by direct sequencing of the mature protein. For these SNPs only six of the eight possible combinations were detected, *GCA* being the most frequent haplotype (0.323), corresponding to Asp[75]-Ile[105] in the mature protein. The new polymorphisms of the ovine *CSN1S2* found in this paper are potential molecular markers that could be used to investigate its association with production traits.

Additional key words: CSN1S2; *Ovis aries;* sequencing; SNPs.

Caseins are encoded by highly polymorphic genes, which genetic variation and functional effects on milk composition and cheesemaking properties have been characterized in cattle and goat species (Martin *et al*., 2002; Caroli *et al*., 2009). Although a variable protein pattern has been also found in the ovine milk (Amigo *et al*., 2000; Chessa *et al*., 2003), the genetic control of the polymorphism was assessed only in a few cases. Thus, SNPs have been described at the ovine α_{S1}-, β- and κ-casein genes (*CSN1S1*, *CSN2* and *CSN3*), which are all expressed at amino acid level (Ferranti *et al*., 1995; Ceriotti *et al*., 2004).

The genetic diversity at mRNA and protein levels is starting to be elucidated for the ovine *CSN1S2* (Boisnard *et al*., 1991; Picariello *et al*., 2009; Giambra & Erhardt, 2012), where eighteen exons have been defined on the basis of comparative analysis among members of the Bovidae family (Rijnkels, 2002). The primary structure of ovine α_{S2}-casein deduced from mRNA sequences (Giambra & Erhardt, 2012) presents six variants to the more common allele A, which are named B (Asp[75] → Tyr and Ile[105] → Val), C (Val[45] → Ile and Ala[48] → Ser), D (Arg[46] → Ser), E (Asn[49] → Asp), F (Asn[200] → Lys) and G (Arg[161] → His).

Although exist high throughput methods for mutation detection, direct sequencing still remains as the most accurate method to determine sequence polymorphisms. The aim of this study was to analyze and characterize the genetic polymorphisms of *CSN1S2* from exons 7 to 11 by sequencing PCR amplicons from conserved sequences. Furthermore, a first approach to the distribution of allele frequencies has been determined in a sample of 17 sheeps from Merino breed.

DNA was isolated from whole blood samples of 17 ewes of Merino breed using a standard commercial kit (*UltraClean™ DNA BloodSpin* kit, MO BIO, Carlsbad, CA, USA). Two sets of primer pairs were designed on the

* Corresponding author: jpadilla@unex.es
Received: 05-07-12. Accepted: 18-01-13.

basis of the available bovine α_{S2}-casein sequence (GenBank M94327) to amplify the exon flanking regions 7-8 and 9-11 of ovine α_{S2}-casein gene. The primer sequences were: 5'-ATAGCAATGCTGAGGGGAAA-3' (E7F) and 5'- ACATCGCAAATAGGGTGAGC-3' (E7R) to amplify the region containing the 7 and 8 exons, and 5'- TTAATGAATTGCCCTTTCTACTCT-3' (E9F) and 5'- GGTGAGGGATGAGTTCAGGA-3' (E9R) to amplify 9, 10 and 11 exon regions.

PCR was carried out in a total volume of 25 µL containing 100 ng of genomic DNA, 200 µM-dNTP, 1.5 mM $MgCl_2$, 125 nM of each primers, 1 U DNA polymerase (DyNAZyme ™ II, Finnzymes, Finland) and 0.8 mg mL^{-1} of BSA (only in E9 reactions). PCR amplification was performed using a PTC-200 thermal cycler (MJ Research, Inc., Watertown, MA, USA) with an initial denaturation step at 95°C for 5 min followed by 30 cycles of 1 min at 95°C for 1 min at 50 or 62°C (E7 and E9 reactions, respectively) and 1 min at 72°C, and a final extension of 15 min at 72°C.

The PCR products were purified with EXOSAP-IT (Amersham Pharmacia Biotech) and sequenced with the forward primers using a fluorescent-labelled dideoxynucleotides termination method, and resolved on an ABI 3130 Genetic Analyser (Applied Biosystems, CA, USA). The SNPs were detected using Sequencing Analysis v5.2 and SeqScape v2.5 softwares (Applied Biosystems), and confirmed by visual examination of the electropherograms. Multiple alignments of the sequences were performed using the ClustalW2/ X (www.clustal.org/clustal2/).

Sequences of the genomic regions covering exons 7-8 and 9-11 of ovine *CSN1S2* have been deposited in the EMBL database, with accession codes FN601396 and FN601350, respectively. Eighteen polymorphic sites (nine transitions and nine transversions) were identified, thirteen located in sequences corresponding to introns 7 (5 SNPs), 8 (2 SNPs), 9 (1 SNP), and 10 (5 SNPs) whereas five are placed in exons 10 (4 SNPs) and 11 (1 SNP). The number of variable sites and the allele frequencies (Table 1) confirm that ovine *CSN1S2* is highly polymorphic, like its bovine and caprine counterparts (Caroli *et al*., 2009). However, none of the SNPs previously identified in exon 7 of other breeds (alleles C, D and E; Boisnard *et al*., 1991; Giambra & Erhardt, 2012) are detected in the Merino population analyzed in this work.

Functional effects of SNPs located at introns of casein genes on production traits have been identified, since mutations within the non-coding sequences have been shown to affect the specific protein expression and, as a consequence, milk composition and cheese-making (Caroli *et al*., 2009). Among the five SNPs identified at intron 10, the adjacent variants g. [435*C* > *A*; 436*T* > *G*] are located at the pyrimidine rich box near the 3' splice acceptor site, which could affect the mRNA processing. In that sense, a SNP at the splice donor sequence of the exon 4 of the bovine *CSN1S1* gene resulted in exon skipping during pre-mRNA processing, with reduced αs1-CN expression (Mohr *et al*., 1994). In addition, two polymorphisms found at the intron 1 splicing enhancer of the equine *CSN2* change the processing pattern of the mRNA (Lenasi *et al*., 2006), and a SNP located at the intron 16 of bovine *CSN1S2* that has been found associated with milk yield (Nilsen *et al*., 2009).

Analysis of deduced amino acid sequences for exons 10 and 11 showed that the three SNPs located at exon 10 g. 143*G* > *A*, g. 155*C* > *T* and g. 176*A* > *C* were synonymous substitutions corresponding to amino acid residues of lysine (p. 71), aspartic (p. 75) and alanine (p. 82), respectively. Among these, g. 155*C* > *T* and g. 176*A* > *C* have been detected by the first time in this work. Beside these, two non-synonymous polymorphisms have been detected in this work, g. 153*G* > *T* and g. *504A* > *G,* affecting the coding sequence for

Table 1. Allele frequencies of polymorphisms at genomic regions spanning exons 7 to 11 and the corresponding introns of ovine CSN1S2 gene

Locus[1]	Allele	Frequency	Allele	Frequency
in7-g. 143A > G	A	0.706	G	0.294
in7-g. 163T > A	T	0.588	A	0.412
in7-g. 226A > C	A	0.676	C	0.323
in7-g. 269A > G	A	0.735	G	0.265
in7-g. 379G > T	G	0.618	T	0.382
in8-g. 545T > C	T	0.294	C	0.706
in8-g. 553C > T	C	0.313	T	0.687
in9-g. 83T > C	T	0.470	C	0.530
ex10-g. 143G > A	G	0.882	A	0.118
ex10-g. 153G > T	G	0.530	T	0.470
ex10-g. 155C > T	C	0.941	T	0.059
ex10-g. 176A > C	A	0.500	C	0.500
in10-g. 202T > A	T	0.500	A	0.500
in10-g. 300G > C	C	0.853	G	0.147
in10-g. 331C > T	T	0.500	C	0.500
in10-g. 435C > A	C	0.735	A	0.265
in10-g. 436T > G	T	0.882	G	0.118
ex11-g. 504A > G	A	0.500	G	0.500

[1] Nucleotide positions are deduced from reference sequences FN601396 and FN601350.

Table 2. Frequencies of the haplotypes in the alleles A and B of ovine CSN1S2 gene

Haplotype	Nucleotide sequence g. 153-155-504[1]	Protein sequence p. 75-105[2]	Frequency	Allele
1	*GCA*	*Asp-Ile*	0.323	A_1
2	*GTA*	*Asp-Ile*	—	A_1
3	*GCG*	*Asp-Val*	0.176	A_2
4	*GTG*	*Asp-Val*	0.029	A_2
5	*TCG*	*Tyr-Val*	0.294	B_1
6	*TTG*	*Tyr-Val*	0.029	B_1
7	*TCA*	*Tyr-Ile*	0.147	B_2
8	*TTA*	*Tyr-Ile*	—	B_2

[1] Positions corresponding to reference sequence FN601350 of ovine CSN1S2 gen. [2] Positions of amino acid in the encoded α_{S2}-casein.

exons 10 and 11, respectively. The polymorphic sites g. 153*G* > *T* (non-synonymous) and g. 155*C* > *T* (synonymous) affect the first and third positions of the same codon, producing the four sequences encoding Asp^{75} (GAC and GAT) or Tyr^{75} (TAC and TAT). The mutation g. 504*A* > *G* at the exon 11 produces change of isoleucine^{105} to valine^{105}. Both non-synonymous amino acid changes, $Asp^{75} \rightarrow$ Tyr and $Ile^{105} \rightarrow$ Val, cause the commonest B allele previously found by direct sequencing of the mature protein in Italian breeds (Chesa *et al*., 2003) or in German breeds, including its mRNA (Giambra *et al*., 2010; Giambra & Erhardt, 2012). Between both polymorphisms, only exchange $Asp^{75} \rightarrow$ Tyr affects the net charge of the protein, which potential effects on milk properties have been previously considered (Picariello *et al*., 2009).

Two new combinations of polymorphic positions, probably recombinants, have been detected by deciphering the haplotypes for non-synonymous variable sites *ex10*- g [153 *G* > *T*; 155 *C* > *T*], and *ex11*- g. 504*A* > *G* (Table 2). Besides the two major alleles A and B of ovine α_{S2}-casein (Asp^{75} and Ile^{105} or Tyr^{75} and Val^{105}, from now A_1 and B_1, respectively), variant forms A_2 and B_2 (Asp^{75} and Val^{105} or Tyr^{75} and Ile^{105}, respectively) occurred with relatively high frequencies. However, among the eight possible haplotypes only six were found, corresponding to alleles A_1 (haplotype *GCA*, in the order g. 153, 155, and 504), A_2 (*GCG* and *GTG*]), B_1 (*TCG* and *TTG*) and B_2 (*TCA*), since *ex10*- g (155 *T*) is a minoritary variant only detected for A2 or B1 haplotypes.

This study has revealed new variants of the ovine CSN1S2 gene, increasing the density of its polymorphism map and, consequently, its potential for genetic analyses. Further studies must be undertaken to reveal their possible association with the production traits.

Acknowledgements

We thank J. Salazar for her assistance in laboratory analysis and to the Spanish Merino Association for providing animal samples. This work was supported by the Junta de Extremadura-FEDER (grant number 2PR03B027).

References

Amigo L, Recio I, Ramos M, 2000. Genetic polymorphism of ovine milk proteins: its influence on technological properties of milk – a review. Int Dairy J 10: 135-149.

Boisnard M, Hue D, Bouniol C, Mercier JC, Gaye P, 1991. Multiple mRNA species code for two non-allelic forms of ovine α_{S2}-casein. Eur J Biochem 201: 633-641.

Caroli AM, Chessa S, Erhardt GJ, 2009. Invited review: milk protein polymorphisms in cattle: effect on animal breeding and human nutrition. J Dairy Sci 92: 5335-5352.

Ceriotti G, Chessa S, Bolla P, Budelli E, Bianchi L, Duranti E, Caroli A, 2004. Single nucleotide polymorphisms in the ovine casein genes detected by polymerase chain reaction-single strand conformation polymorphism. J Dairy Sci 87: 2606-2613.

Chessa S, Bolla P, Dario C, Pieragostini E, Caroli A, 2003. Genetic milk protein polymorphisms in the Gentile di Puglia ovine breed: monitoring by isoelectric focusing. Sci Tecn Latt-Cas 54: 191-198.

Ferranti P, Malorni A, Nitti G, Laezza P, Pizzano R, Chianese L, Addeo F, 1995. Primary structure of ovine α_{S1}-caseins: localization of phosphorylation sites and characterization of genetic variants A, C and D. J Dairy Res 62: 281-296.

Giambra IJ, Erhardt G, 2012. Molecular genetic characterization of ovine CSN1S2 variants C and D reveal further important variability within CSN1S2. Anim Genet 43: 642-645.

Giambra IJ, Jäger S, Erhardt G, 2010. Isoelectric focusing reveals additional casein variants in German sheep breeds. Small Ruminant Res 90: 11-17.

Lenasi T, Peterlin BM, Dovc P, 2006. Distal regulation of alternative splicing by splicing enhancer in equine β-casein intron 1. RNA 12: 498-507.

Martin P, Szymanowska M, ZwIerzchowski L, Leroux C, 2002. The impact of genetic polymorphisms on the protein composition of ruminant milks. Reprod Nutr Dev 42: 433-459.

Mohr U, Koczan D, Linder D, Hobom G, Erhardt G, 1994. A single point mutation results in A allele-specific

exon skipping in the bovine α_{S1}-casein mRNA. Gene 143: 187-192.

Nilsen H, Olsen HG, Hayes B, Sehested E, Svendsen M, Nome T, Meuwissen T, Lien S, 2009. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. Genet Sel Evol 41: 24.

Picariello G, Rignanese D, Chessa S, Ceriotti G, Trani A, Caroli A, Di Luccia A, 2009. Characterization and genetic study of the ovine α_{S2}-Casein (CSN1S2) Allele B. Protein J 28: 333-340.

Rijnkels M, 2002. Multispecies comparison of the casein gene loci and evolution of casein gene family. J. Mammary Gland Biol 7: 327-345.