J. Dairy Sci. 93:792–795
doi:10.3168/jds.2009-2615

Short communication: Molecular genetic characterization of ovine α_{S1}-casein allele H caused by alternative splicing

I. J. Giambra,* L. Chianese,† P. Ferranti,† and G. Erhardt*[1]
*Department of Animal Breeding and Genetics, Justus-Liebig-University, 35390 Gießen, Germany
†Dipartimento di Scienza degli Alimenti, University of Naples "Federico II," 80055 Portici, Naples, Italy

ABSTRACT

Sequencing of ovine *CSN1S1*H* cDNA showed an absence of exon 8 in comparison with GenBank sequences; the absence was confirmed by protein sequencing. We demonstrated that this allelic aberration is the result of a deletion of 4 nucleotides, the last 3 of exon 8 and the first 1 of intron 8, which are replaced by an insertion of 13 nucleotides in the DNA sequence. The insertion is a precise duplication of a part of the adjacent intronic sequence of *CSN1S1*C″*. These sequence differences result in an inactivation of the splice donor sequence distal to exon 8, leading to upstream exon skipping during the serial splice reactions of the ovine *CSN1S1*H* pre-mRNA, and may affect the specific casein expression as well as protein characteristics.
Key words: ovine, *CSN1S1*, messenger ribonucleic acid, exon skipping

Among the ovine CN, the 4 fractions α_{S1}-, α_{S2}-, β-, and κ-CN, encoded by the genes *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3*, respectively, are distinguished. Milk proteins and their genetic polymorphisms are of interest in ruminant breeding; several studies have revealed their effects on production traits (Martin et al., 2002; Boettcher et al., 2004) and on cheese-making properties (Ikonen et al., 1999; Pirisi et al., 1999; Hallén et al., 2007). Furthermore, milk protein variants and their allelic or phenotypic frequencies can be used for the molecular tracing of typical cheeses (Chianese et al., 2009) and as population genetic markers (Mahé et al., 1999; Ibeagha-Awemu et al., 2007).

So far, 7 phenotypes (A, B, C, D, E, F, and X) of α_{S1}-CN in ovine milk have been identified by protein electrophoresis (Chianese et al., 1996; Pirisi et al., 1999; Wessels et al., 2004), whereas primary structures have been determined only for α_{S1}-CN A, C, D, and E (Ferranti et al., 1995; Chianese et al., 2007). Within *CSN1S1*C*, Ceriotti et al. (2005) distinguished between 2 haplotypes, *C′* and *C″*, at the DNA level. Recently, the α_{S1}-CN phenotype X was recovered by isoelectric focusing (**IEF**) in East Friesian dairy sheep and named *CSN1S1*H* (Giambra et al., 2009).

Therefore, the aims of this study were to characterize *CSN1S1*H* at the protein, mRNA, and DNA levels to better understand *CSN1S1* variants in sheep, and to develop a DNA-based test for their identification.

Milk samples of 4 East Friesian dairy sheep with *CSN1S1* genotypes *HH* (n = 1) and *CH* (n = 3), previously typed by IEF, were used for molecular genetic analysis. For comparative analysis, 6 milk and blood samples of reference sheep with known *CSN1S1* genotypes (*AC″*: n = 1; *C″C″*: n = 2; *C″D*: n = 2; *DD*: n = 1), typed by IEF and by DNA sequencing, were used.

From each of the 10 samples, about 10 mL of fresh milk were collected and centrifuged (10 min, 2,200 × *g*) and the resulting pellet was washed according to Boutinaud et al. (2002). Total RNA was extracted with the Invisorb Spin Plant RNA mini kit (Invitek, Berlin, Germany). The RNA was assembled immediately for cDNA synthesis in 4 separate reactions for each sample with the Verso cDNA kit using an Oligo-dT primer (both from Thermo Fisher Scientific, Waltham, MA).

The resulting cDNA was used as PCR template for each animal using 2 of the 4 cDNA generated per mRNA sample in 2 separate PCR. The forward primer (5′-cttcttcccagtcttgggttc-3′) is located in the first exon of *CSN1S1* mRNA, whereas the reverse primer (5′-aattttcaatcccatcaaagacc-3′) spans exons 18 and 19. Amplified PCR fragments were cloned with pGEM-T and pGEM-T Easy Vectors (Promega, Mannheim, Germany) according to Shin et al. (2005). Fifty recombinant clones per animal were sequenced bidirectionally using an ABI Prism 377 DNA analyzer, Big Dye Terminator sequencing kit v1.1, and vector primers M13-Forward and M13-Reverse (all from Applied Biosystems, Foster City, CA).

Genomic DNA was isolated from blood according to Montgomery and Sise (1990) and from milk somatic cells according to Lühken et al. (2009) and was used for 4 PCR (Table 1; PCR 1–4) concerning DNA sequence from exon 7 to exon 9 in overlapping fragments. Bi-

Received August 5, 2009.
Accepted October 30, 2009.
[1]Corresponding author: Georg.Erhardt@agrar.uni-giessen.de

directional direct sequencing of the 4 PCR fragments (1–4) was done as described above, using the same primers used for PCR.

To establish a DNA-based test, the forward primer used in PCR 5 (Table 1) was labeled with indodicarbocyanine (Cy5) to enable fragment length analysis using an A.L.F.express sequencer and AlleleLocator 1.03 software (both from Amersham Biosciences, Freiburg, Germany) for all carriers of *CSN1S1*H* and the reference samples. The PCR products were run on a 5.5% Long Ranger gel (0.5 mm, 6 *M* urea) at 800 V, 60 mA, and 49°C for 280 min using internal (202 bp and 266 bp) and external (266 bp ladder) size markers. The PCR 5 product was also used for RFLP analysis with the restriction enzyme *DdeI* for identification of *CSN1S1*H*.

To supplement mRNA and DNA sequencing results, protein isolation and mass spectrometry analysis were performed simultaneously for the *CSN1S1 C″C″*, *CH*, and *HH* samples according to Chianese et al. (2009).

Bioinformatic analysis of nucleotide and AA sequences were performed with the programs ChromasPro 1.32 (Technelysium Pty Ltd., Queensland, Australia), DNASTAR (DNAstar Inc., Madison, WI), and NNSPLICE 0.9 (Reese et al., 1997).

Nucleotide sequences of cloned PCR fragments revealed that exon 8 was missing in *CSN1S1*H* cDNA/mRNA (GenBank acc. no. FJ440846) compared with GenBank acc. no. NM_001009795 and GenBank acc. no. FJ440845, which is the full-length coding sequence of *CSN1S1*C″*, containing exon 16, established in this study for the first time. The lack of exon 8 coding for 8 AA (AA 51–58; DQAMEDAK), which does not influence the reading frame, resulted in a truncated protein. Mass spectrometry analysis of HPLC-purified dephosphorylated H variant provided a molecular mass of 21,872 Da, which indicated the presence of a deleted variant of α_{S1}-CN that lacked the sequence 51 to 58; this was further confirmed by tandem mass spectrometry analysis of tryptic digest of the variant protein (data not shown).

Analysis of the genomic ovine *CSN1S1* sequence from exon 7 to exon 9 showed that the sequence corresponding to exon 8 is present in the DNA sequence. The alignment of *CSN1S1*H* with *CSN1S1*A*, *C″*, and *D* DNA sequences showed a deletion of 4 nucleotides (the last triplet of exon 8 and the first nucleotide of intron 8) and an insertion of 13 bp (g.739_742delAAGGinsT TATTTTAATAAA) in *CSN1S1*H* (Figure 1). This insertion is a precise duplication of the adjacent intronic sequence (g.751–763) of *CSN1S1*C″* (GenBank acc. no. FJ440847). These changes delete the 5′-donor splice site of intron 8 with its special AG/GT-configuration (g.740–747; GenBank acc. no. FJ440847) in *CSN1S1*H*

Table 1. Primer sequences, reaction conditions, and resulting product sizes used for characterization of ovine *CSN1S1* DNA sequence from exon 7 to exon 9

PCR number	PCR name[1]	Primer forward, 5′– 3′ (GenBank acc. no. and location)	Primer reverse, 5′–3′ (GenBank acc. no. and location)	Mg concentration (m*M*)	AT[2] (°C)	Cycles (n)	Product size (bp)
1	Ex7F-In8R	gggagtgaatcaattgaggtaag (NM_001009795; r.239–256 in combination with X59856; g.16685–16689)	caattgcagtggatagctttctt (X59856; g.7487–17465)	1.0	60	45	821
2	In7F-In8bR	ttctctattgccacccatttcta (FJ440847; g.603–625)	actgactgaatggacgtgagttt (X59856; g.17904–17882)	3.0	59	35	643
3	In8F-In8cR	gctatgctgttgttcagtcactc (FJ440847; g.1126–1148)	gaacaatcaatttcagaatggaa (X59856; g.18478–18456)	1.0	56	35	699
4	In8bF-In9R	cccaagagatgactatttttgctg (FJ440847; g.1639–1661)	tgttttcttacattttgggggttg (X59856; g.19081–19103)	1.0	53	35	805
5	In7FCy5-In8dR	atttccttggcatccatttttat (FJ440847; g.625–646)	tagcactgcttgaggagttcaat (FJ440847; g.855–832)	2.0	60	35	231

[1]Ex = exon; In = intron.
[2]AT = annealing temperature.

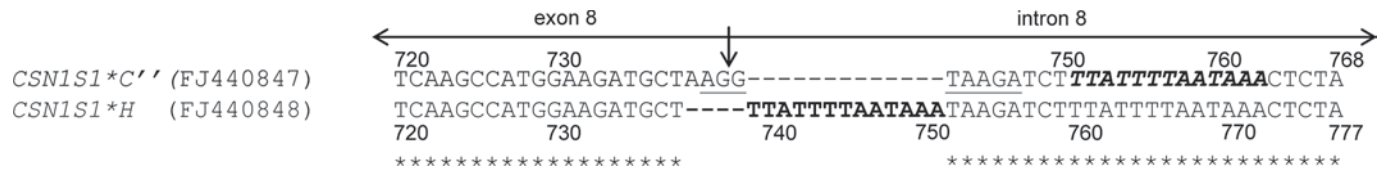

Figure 1. Genomic DNA sequence alignment of ovine *CSN1S1* variants *C″* (nucleotides 720–768; GenBank acc. no. FJ440847) and *H* (nucleotides 720–777; GenBank acc. no. FJ440848). Stars mark identical nucleotides. Blanks within the sequence of *CSN1S1*C″* mark a region where an insertion in *CSN1S1*H* occurred. Blanks in *CSN1S1*H* mark a deletion of the last 3 nucleotides of exon 8 and the first nucleotides of intron 8. Bold letters mark an insertion in *CSN1S1*H* corresponding to intron 8 nucleotides 751–763 in *CSN1S1*C″*. Underlined nucleotides display the 5′-donor splice site of intron 8 in *CSN1S1*C″*. The arrow indicates the junction of exon 8 and intron 8.

as shown in Figure 1. In the junction of exon 8 and intron 8 in *CSN1S1*H*, no splice site could be identified in silico by NNSPLICE 0.9 (Reese et al., 1997). The sequence differences in ovine *CSN1S1*H* at the 5′-splice site alter the U1 small nuclear RNA recognition sequence and inhibit accurate splice site selection and the splicing process.

The characterization of *CSN1S1*H* with a deletion in the exon and an insertion in the intron DNA sequence describes an ovine CN variant caused by exon skipping, which is different in comparison with *CSN1S1* in cattle (Mohr et al., 1994) and goats (Leroux et al., 1992; Grosclaude and Martin, 1997), where SNP in the splice donor sequence of the following intron are responsible for alternative splicing.

Fragment length analysis of PCR 5 revealed sequence length differences of 9 nucleotides between *CSN1S1*H* and *C″* (nonallele H: 231 bp; allele H: 240 bp; Figure 2). These results were confirmed by PCR-RFLP analyses, whereas restriction of PCR 5 product with *DdeI* resulted in a control fragment of 89 bp in all cases, 2 additional fragments in samples without *CSN1S1* allele *H* (24 bp and 118 bp), and only 1 additional fragment in carriers of *CSN1S1*H* (151 bp; Figure 3). Fragment length and PCR-RFLP analyses can be used as DNA tests for the identification of *CSN1S1*H* in animals independent from age, sex, and lactation.

The widespread missing of short exon sequences in different CN and in different species could be supported by the extremely split architecture characteristic of *CSN1S1* and *CSN1S2* in various species with small exons in the coding region (Rijnkels, 2002), caused by evolutionary inter- and intragenic duplications (Groenen et al., 1993).

The possible effects of *CSN1S1*H* on milk composition, milk yield, and cheese-making abilities are the subject of ongoing studies. This is interesting because of the correlation between abbreviated caprine *CSN1S1*F* and *G* and reduced protein content (Leroux et al., 1992; Grosclaude and Martin, 1997) and the correlation between bovine *CSN1S1*A* and lower cheese-making ability and cheese quality (Ng-Kwai-Hang, 2006). Quantitative analysis of the expressed protein by integration of mass ion peaks of dephosphorylated CN resulted in a 74% lower expression level for α_{S1}-CN *H* compared with the reference *C″* variant (data not shown). This is in agreement with the results of Wessels et al. (2004), who found significantly lower CN content in milk from sheep

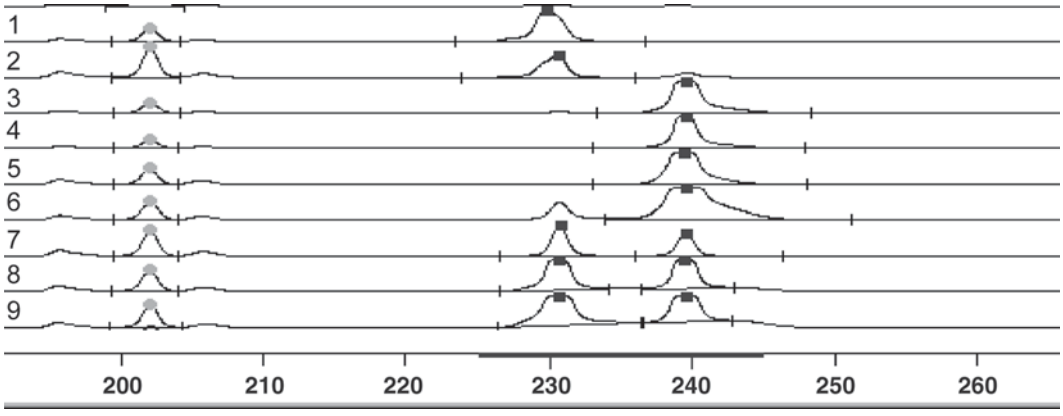

Figure 2. Fragment length analysis of ovine *CSN1S1* using an A.L.F.express sequencer and AlleleLocator 1.03 software (both Amersham Biosciences, Freiburg, Germany). Peak at position of 202 bp indicates external marker. Peak at position of 231 bp marks non-*CSN1S1*H*, whereas peak at position of 240 bp marks *CSN1S1*H*. Lanes 1 and 2: non-*CSN1S1*H*; lanes 3, 4, 5, and 6: *CSN1S1 HH*; lanes 7, 8, and 9: *CSN1S1*H* carrier.

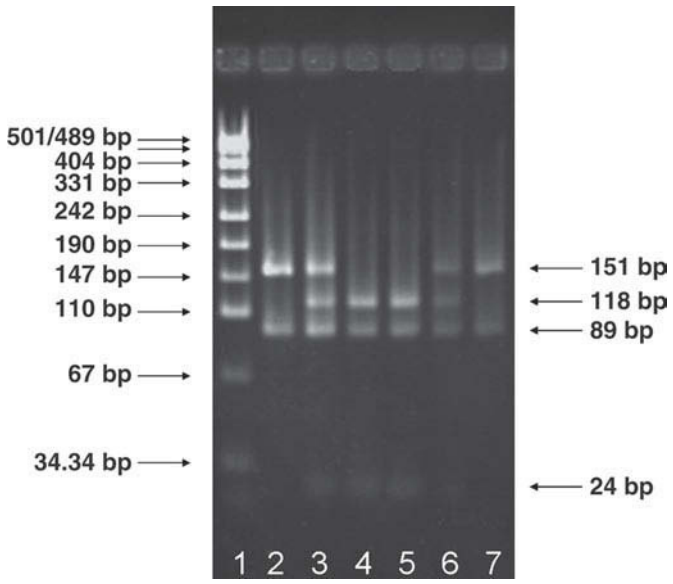

Figure 3. Separation of *Dde*I-digested PCR products in a 3.5% agarose gel stained with ethidium bromide. Lane 1: marker; lanes 2, 7: *CSN1S1 HH*; lanes 3, 6: *CSN1S1 C″H*; lanes 4, 5: *CSN1S1 C″C″*.

with *CSN1S1 CX*, which is identical to the *CSN1S1* genotype *CH*, compared with *CSN1S1 CC* sheep.

ACKNOWLEDGMENTS

We thank Gesine Lühken (Department of Animal Breeding and Genetics, Justus-Liebig-University Gießen, Germany) and the anonymous reviewers for valuable comments.

REFERENCES

Boettcher, P. J., A. Caroli, A. Stella, S. Chessa, E. Budelli, F. Canavesi, S. Ghiroldi, and G. Pagnacco. 2004. Effects of casein haplotypes on milk production traits in Italian Holstein and Brown Swiss cattle. J. Dairy Sci. 87:4311–4317.

Boutinaud, M., H. Rulquin, D. H. Keisler, J. Djiane, and H. Jammes. 2002. Use of somatic cells from goat milk for dynamic studies of gene expression in the mammary gland. J. Anim. Sci. 80:1258–1269.

Ceriotti, G., F. Chiatti, R. Bolla, M. Martini, and A. Caroli. 2005. Genetic variability of the ovine alpha(S1)-casein. Ital. J. Anim. Sci. 4:64–66.

Chianese, L., S. Caira, G. Garro, and F. Addeo. 2007. Primary structure of ovine deleted variant α_{S1}-CN E. Page 70 in Proc. 5th International Symposium on the Challenge to Sheep and Goats Milk Sectors, Alghero–Sardinia. Italy. International Dairy Federation, Brussels, Belgium.

Chianese, L., G. Garro, R. Mauriello, P. Laezza, P. Ferranti, and F. Addeo. 1996. Occurrence of five alpha s1-casein variants in ovine milk. J. Dairy Res. 63:49–59.

Chianese, L., M. Quarto, F. Pizzolongo, M. G. Calabrese, S. Caira, R. Mauriello, S. De Pascale, and F. Addeo. 2009. Occurrence of genetic polymorphism at the alpha(s1)-casein locus in Mediterranean water buffalo milk. Int. Dairy J. 19:181–189.

Ferranti, P., A. Malorni, G. Nitti, P. Laezza, R. Pizzano, L. Chianese, and F. Addeo. 1995. Primary structure of ovine alpha s1-caseins: Localization of phosphorylation sites and characterization of genetic variants A, C and D. J. Dairy Res. 62:281–296.

Giambra, I. J., S. Jäger, and G. Erhardt. 2009. Isoelectric focusing reveals additional casein variants in German sheep breeds. Small Rumin. Res. (accepted)

Groenen, M. A., R. J. Dijkhof, A. J. Verstege, and J. J. Van Der Poel. 1993. The complete sequence of the gene encoding bovine alpha s2-casein. Gene 123:187–193.

Grosclaude, F., and P. Martin. 1997. Casein polymorphisms in the goat. Pages 241–253 in Milk Protein Polymorphism. International Dairy Federation, Brussels, Belgium.

Hallén, E., T. Allmere, J. Naslund, A. Andren, and A. Lunden. 2007. Effect of genetic polymorphism of milk proteins on rheology of chymosin-induced milk gels. Int. Dairy J. 17:791–799.

Ibeagha-Awemu, E. M., E. M. Prinzenberg, O. C. Jann, G. Lühken, A. E. Ibeagha, X. Zhao, and G. Erhardt. 2007. Molecular characterization of bovine CSN1S2*B and extensive distribution of zebu-specific milk protein alleles in European cattle. J. Dairy Sci. 90:3522–3529.

Ikonen, T., K. Ahlfors, R. Kempe, M. Ojala, and O. Ruottinen. 1999. Genetic parameters for the milk coagulation properties and prevalence of noncoagulating milk in Finnish dairy cows. J. Dairy Sci. 82:205–214.

Leroux, C., N. Mazure, and P. Martin. 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. J. Biol. Chem. 267:6147–6157.

Lühken, G., A. Caroli, E. M. Ibeagha-Awemu, and G. Erhardt. 2009. Characterization and genetic analysis of bovine α_{s1}-casein *I* variant. Anim. Genet. 40:479–485.

Mahé, M. F., G. Miranda, R. Queval, A. Bado, P. S. Zafindrajaona, and F. Grosclaude. 1999. Genetic polymorphism of milk proteins in African *Bos taurus* and *Bos indicus* populations. Characterization of variants alpha(s1)-Cn H and kappa-Cn J. Genet. Sel. Evol. 31:239–253.

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

Mohr, U., D. Koczan, D. Linder, G. Hobom, and G. Erhardt. 1994. A single point mutation results in A allele-specific exon skipping in the bovine alpha s1-casein mRNA. Gene 143:187–192.

Montgomery, G. W., and J. A. Sise. 1990. Extraction of DNA from sheep white blood cells. N.Z. J. Agric. Res. 33:437–441.

Ng-Kwai-Hang, K. F. 2006. Genetic variants of milk proteins and their effects on the yield and quality of cheese. CAB Reviews 56:1–11. doi:10.1079/PAVSNNR20061056

Pirisi, A., G. Piredda, C. M. Papoff, R. Di Salvo, S. Pintus, G. Garro, P. Ferranti, and L. Chianese. 1999. Effects of sheep alpha s1-casein CC, CD and DD genotypes on milk composition and cheesemaking properties. J. Dairy Sci. 66:409–419.

Reese, M. G., F. H. Eeckman, D. Kulp, and D. Haussler. 1997. Improved splice site detection in Genie. J. Comput. Biol. 4:311–323.

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

Shin, H. J., H. Y. Park, S. J. Jeong, H. W. Park, Y. K. Kim, S. H. Cho, Y. Y. Kim, M. L. Cho, H. Y. Kim, K. U. Min, and C. W. Lee. 2005. STAT4 expression in human T cells is regulated by DNA methylation but not by promoter polymorphism. J. Immunol. 175:7143–7150.

Wessels, G., H. Hamann, G. Erhardt, and O. Distl. 2004. Genotypeffekte von Milchprotein-Polymorphismen auf die Milchleistung beim Ostfriesischen Milchschaf. Berl. Munch. Tierarztl. Wochenschr. 117:414–419.