Journal of Dairy Research (2005) **72** 317–321. © *Proprietors of Journal of Dairy Research 2005*
doi:10.1017/S0022029905000932 *Printed in the United Kingdom*

A single nucleotide polymorphism in the sheep κ-casein coding region

Maria Feligini[1]*, Slavica Vlaco[1], Vlatka Cubric Curik[2], Pietro Parma[1], GianFranco Greppi[1,3] and Giuseppe Enne[1,4]

[1] Istituto Sperimentale Italiano ''Lazzaro Spallanzani'', Laboratorio di Epigenomica Applicata, Viale Giovanni XXIII 7, I-26900 Lodi, Italy
[2] Dairy Science Department, Faculty of Agriculture, University of Zagreb, Svetosimunska 25, HR-10 000 Zagreb, Croatia
[3] Dipartimento di Clinica Medica Veterinaria, Università degli Studi di Milano, Via Celoria 10, Milan, Italy
[4] Dipartimento di Scienze Zootecniche, Facoltà di Agraria, Università di Sassari, Via De Nicola, Sassari, Italy

Received 27 July 2004 and accepted for publication 22 November 2004

Genetic polymorphisms in *CSN3* gene in Pag (Croatia), Sarda (Italy) and Pramenka (Serbia) sheep breeds were investigated. A single nucleotide polymorphism (SNP) was localized by sequence analysis (sequence submitted to GenBank under accession AY237637) relying on an original primer pair. Primers for sequencing (κ-casF and κ-casR) were designed on the available *CSN3* sequences to amplify the genomic region encoding the major part of the mature protein (exon 4). An SNP was detected at position 237 of the sheep κ-casein mRNA (reference sequence: GenBank X51822), where a thymine was substituted for a cytosine. The SNP was typed by conventional PCR and SYBR Green I-based real-time PCR. C and T alleles were discriminated using a dedicated set of primers that consisted of one common forward primer (SNP-TC) and two reverse primers (SNP-T and SNP-C), the latter two differing in the 3′ end base and in the presence of a 12 bp poly-G tail in SNP-C. The SNP was found in both the heterozygous and the homozygous state in Sarda and Pramenka breeds, and in the heterozygous state only in the Pag breed. The observed allelic frequencies of the SNP were 0·12 in Pag, 0·27 in Sarda and 0·45 in Pramenka.

Keywords: *CSN3* gene, ASA-PCR, real-time PCR.

Caseins are a family of phosphoproteins synthesized in the mammary gland in response to lactogenic hormones and secreted as micelles. The importance of caseins to the dairy industry has made them a popular target for genetic and biochemical studies (Ginger & Grigor, 1999). Among the four caseins (α_{s1}-, α_{s2}-, β- and κ-casein) κ-casein is the only one that is soluble in the presence of calcium ions and is glycosylated (Mercier & Vilotte, 1993). κ-Casein is located mainly on the surface of the casein micelle and it contributes to determining the size of the micelle (Waugh, 1971). Analysis of the κ-casein genomic region in the species where this feature is described reveals a conserved exon-intron structure (Baranyi et al. 1996; Edlund et al. 1996). Exon 1 is a short exon of the 5′ untraslated sequence. Exon 2 contains the rest of the untraslater 5′ region and 19 codons (of 21) specifying the signal peptide; the rest of this domain and the first amino-acids of the mature peptide are encoded by exon 3 (33 bp long). Finally the major part of the mature peptide is encoded by exon 4 (507 bp long) and the 3′ untranslated sequence is encoded by the last portion of exon 4 and the whole exon 5 (165 bp long). Lengths of exons 4 and 5 were inferred from available goat (D14371, D14372, D14373), rabbit (U44058) and human (NM_005212) sequences. In sheep the κ-casein gene was previously mapped to chromosome region 6q23-31 (Ansari et al. 1994) and between marker Mc0140 and BMS2460 at 98·30 cM of chromosome 6 in SM3 reference map (Maddox et al. 2001).

κ-Casein polymorphism is well described in cattle (Grosclaude, 1988; Pérez-Rodríguez et al. 1998; Prinzenberg et al. 1999; Damiani et al. 2000; Soria et al. 2003), goat (Russo et al. 1986; Di Luccia et al. 1990; Martin, 1993; Vegarud et al. 1999; Caroli et al. 2001; Feligini et al. 2002; Jann et al. 2004) and sheep (Ordás et al. 1997; Ceriotti et al. 2004). In cattle, a positive association between some genetic variants of *CSN3* and milk characteristics of economical importance has been reported (see Ng-Kwai-Hang, 1998). Nevertheless, until now no polymorphic form has been reported for the sheep κ-casein

*For correspondence; e-mail: maria.feligini@isils.it

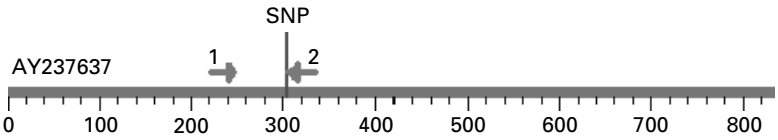

Fig. 1. Position of the genotyped locus within the sequenced region (*CSN3* exon 4, GenBank accession n. AY237637). 1: forward primer SNP-TC (positions 251–272); 2: reverse primers SNP-T and SNP-C (positions 305–324 and 305–324+poly-G tail respectively). SNP is located at position 305, corresponding to position 237 in mRNA sequence (GenBank accession n. X51822).

except for the evidence of *Hind*III, *Pvu*II and *Pst*I polymorphic patterns by using a bovine (Di Gregorio et al. 1991) and ovine (Leveziel et al. 1991) κ-casein cDNA probe in a Southern blot experiment. In sheep, polymorphism in *CSN3* has been reported in an Italian breed (Ceriotti et al. 2004).

Here we focus on detecting polymorphisms in two Balcan breeds (Pag and Pramenka) raised mainly for milk production and in Sarda breed by means of new allele specific polymerase chain reaction (ASA-PCR) and real-time polymerase chain reaction (SG-ASA-PCR) protocols.

Materials and Methods

DNA samples

A total of 120 DNA samples from unrelated Pag sheep (Croatia; $n=36$), Pramenka (Serbia; $n=40$) and Sarda (Italy; $n=44$), were used for the molecular analyses. DNA was extracted from whole blood (Sarda and Pag) and individual milk samples (Pramenka) using a standard commercial kit (Quiagen blood kit, Valence CA, USA; NucleoSpin Food, Macherey-Nagel, Düren, Germany).

Sequencing: primer design, PCR reaction

Primers were designed on the basis of the available sheep κ-casein sequence (GenBank accession number AF165792 and X51822) to amplify the genomic region encoding the major part of the mature peptide (exon 4). The primer sequences were the following: 5′-ttcactccagcctacaatacca-3′ (K-casF) and 5′-ttgctcatttacctgcgttg-3′ (K-casR).

PCR was carried out in a total volume of 50 µl containing 100 ng of genomic DNA, 200 µM-dNTP, 10 mM-Tris–HCl, pH 8·3, 50 mM-KCl, 1·5 mM-$MgCl_2$, 100 nM-primers, 1·5 U DNA polymerase (AmpliTaq Gold, Perkin Elmer). PCR amplification was performed using a GeneAmp 9600 thermal cycler (Perkin Elmer), with an initial 95 °C denaturation of 10 min followed by 30 cycles of 15 s at 95 °C, 30 s at 58 °C, 1 min at 72 °C and final extension of 10 min at 72 °C. The size of the expected product was 583 bp (GenBank accession number AY237637, positions 180–762). The PCR product was purified using a silica gel column (Ultra Clean PCR Clean-Up, MoBio, Solana Beach CA, USA) sequenced with forward primers using a fluorescent-labelled dideoxy-nucleotides termination method, and resolved on an Applied Biosystems 3100 ABI PRISM automated DNA sequencer. Sequencing was performed on 80 samples (36 Pag and 44 Sarda).

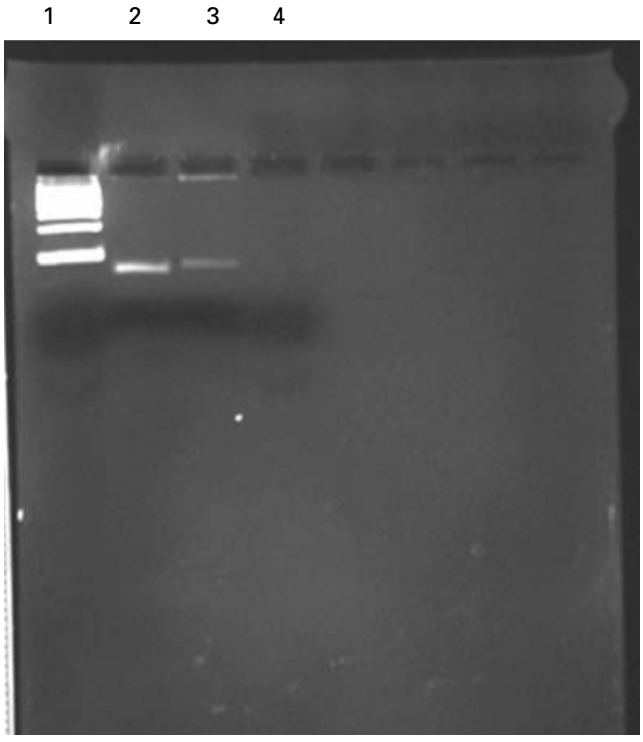

Fig. 2. Conventional ASA-PCR showing the separate amplifications of a single heterozygous sample (CT) in 3% agarose gel. Lane 2: T allele (85 bp); lane 3: C allele (97 bp). Lane 1: 100 bp DNA Ladder (BioLabs). Lane 4: double distilled water.

ASA-PCR Primers

Genotyping of the identified SNP was performed using a set of primers combining one forward primer (SNP-TC; 5′-cttcgatgacaaaatagccaa-3′) and two reverse primers (SNP-T; 5′-aattgagtccataactagga-3′ and SNP-C; 5′-gggg-gggggggaattgagtccataactaggg-3′). Our optimized procedure involved two reactions per sample containing SNP-TC/SNP-T and SNP-TC/SNP-C primer pairs respectively. A poly-G tail was added to the SNP-C reverse in order to introduce a detectable difference in the melting curve profiles opf the two variants; this also allowed the PCR products to be distinguished on agarose gel.

https://doi.org/10.1017/S0022029905000932 Published online by Cambridge University Press

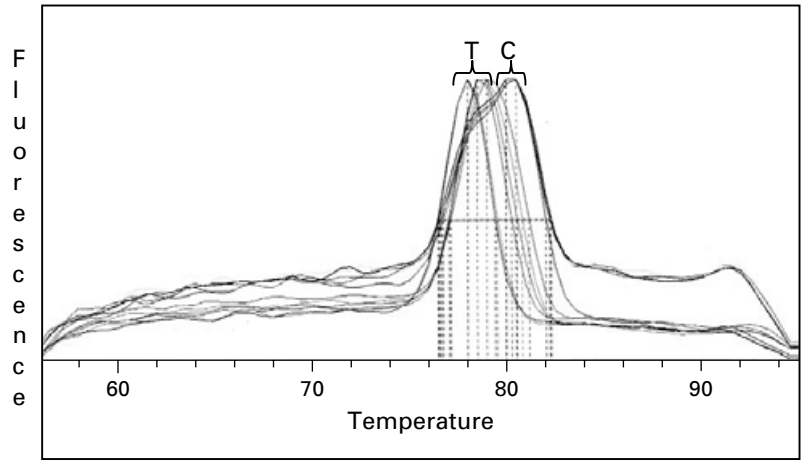

Fig. 3. Real-time PCR genotyping of six homozygous samples (three TT genotypes and three CC genotypes). T and C alleles are discriminated on the basis of melting curve (fluorescence *v.* temperature plots); different characteristics are determined by the differences in the sequence of the corresponding reverse primer.

Allelic Specific Amplification (*ASA-PCR*)

Conventional PCR amplification was performed in a 50-μl reaction volume containing 50 ng DNA, 200 μM-dNTP, 10 mM-Tris–HCl, pH 8·3, 50 mM-KCl, 1·5 mM-$MgCl_2$, 100 nM-primers, 1·5 U DNA polymerase (AmpliTaq Gold, Perkin Elmer). The PCR reaction was carried out in a GeneAmp 9600 thermal cycler (Perkin Elmer) with an initial 95 °C denaturation of 10 min followed by 35 cycles of 30 s at 95 °C, 30 s at 56 °C, 30 s at 72 °C and final extension of 5 min at 72 °C. The PCR product, 85 bp for the T variant and 97 bp for the C variant, was electrophoresed in a 3% agarose gel stained with ethidium bromide. Electrophoretic runs were conducted at constant 100 V for 20 min using TBE as running buffer.

Real-time PCR (*SG-ASA-PCR*)

Real-time PCR amplification was performed in a 30-μl reaction volume containing: 20–30 ng DNA, 200 μM-dNTP, 10 mM-Tris–HCl at pH 8·3, 50 mM-KCl, 1·5 mM-$MgCl_2$, 100 nM-primers, 1 : 50 000 SYBR Green I (Stratagene, La Jolla CA, USA), 0·9 U of DNA Polymerase (AmpliTaq Gold, Perkin Elmer) PCR amplification was carried out using a PTC-200 Opticon™ thermocycler (MJ Research, Reno NV, USA). Real-time PCR temperature cycling conditions were the same as for conventional PCR except that the number of cycles was raised to 40. The specificity of the reaction was monitored by determining the melting temperature of the product (Wittwer et al. 2001).

Two reactions per sample, along with one no-template negative control (water control), were prepared to determine the characteristic melting curve profiles on the basis of the SNP identified. Melting curves were determined from 55 to 95 °C at 0·01 deg C/s.

Amplicons were also analysed in agarose gel to verify that the expected fragment was produced and to confirm the correlation between the gel and fluorescence data.

Results and Discussion

An SNP was detected at position 237 of the sheep κ-casein mRNA (GenBank accession number X51822) where a thymine was substituted for a cytosine whilst the corresponding amino acid remains a tyrosine. We completely sequenced the *CSN3* exon 4 locus from Pag, Pramenka and Sarda breeds; the sequence was submitted to GenBank database under the accession number AY237637.

The frequency of the SNP was investigated in 120 unrelated sheep representing all three breeds by conventional Allelic Specific Amplification PCR (Parsons & Heflich, 1997), followed by agarose gel electrophoresis, and real-time PCR (Rajeevan et al. 2001; Wittwer et al. 2001). A primer set was designed to amplify the relevant region of the two alleles (Fig. 1). While the forward one was common, the two reverse primers differed in two features: the 3′ end, which bonds to the template's varying base, and a 12 bp-long guanine polymer attached to the C-variant reverse 5′ end. The latter had the double purpose of increasing the molecular weight of the conventional PCR product, affecting electrophoresis, and shifting up melting temperature in real-time assays, thus changing the melting curve characteristics. Owing to the presence of this poly-G tail, different length products were synthesized for the two alleles (85 bp for T allele and 97 bp for C allele). This design allowed us to set up separate reactions for T and C variants, so that we could deduce the genotype of the samples by combining pairs of positive/negative results (Fig. 2). In real-time PCR assays, genotyping was achieved by means of melting curve analysis, which proved capable of directly identifying T and C alleles from their melting profiles (Fig. 3). The melting temperature of the specific

https://doi.org/10.1017/S0022029905000932 Published online by Cambridge University Press

Table 1. Genotype distribution and allele frequencies of the κ-casein polymorphism in the ovine breeds typed

Origin	Breeds	Genotypes			Allele frequency	
		TT	TC	CC	T	C
Croatia	Pag	0	9	27	0·12	0·88
Italy	Sarda	6	12	26	0·27	0·73
Serbia	Pramenka	6	24	10	0·45	0·55

product and that of the primer dimer were separated by at least 3–4 °C. This real-time PCR technique allows fast and accurate allele detection, without the use of toxic compounds like ethidium bromide; it allows genotyping up to 96 samples at the same time in a large DNA concentration range, with a high identification accuracy (Sanders Sevall, 2001).

The resulting genotype distribution is reported in Table 1. The highest frequency was observed in Pramenka (0·45), whilst the lowest was observed in Pag (0·12). We observed this SNP also in the Sarda breed, where it had an allelic frequency of 0·27, while it was formerly detected only once in a heterozygous animal belonging to an Italian breed as reported by Ceriotti et al. (2004).

Owing to the influence of *CSN3* gene mutations on cheesemaking as previously reported for bovine species (Aleandri et al. 1990; Ng-Kwai-Hang et al. 1990; Rahali & Ménard, 1991; Macheboeuf et al. 1993; Lodes et al. 1996), their identification in ovine milk is even more important because sheep milk is commonly used solely for cheesemaking.

This research was funded by a grant from FIRB project (Prof Gian Franco Greppi). We thank Dr Ivan Bonizzi for his assistance in laboratory analysis.

References

Aleandri R, Buttazzoni LG, Schneider JC, Caroli A & Davoli R 1990 The effects of milk protein polymorphisms on milk components and cheese-producing ability. *Journal of Dairy Science* **73** 241–255

Ansari HA, Pearce PD, Maher DW, Malcolm AA, Wood NJ, Phua SH & Broad TE 1994 Regional mapping of loci from human chromosome 2q to sheep chromosome 2q. *Genomics* **20** 122–124

Baranyi M, Aszodi A, Devinoy E, Fontaine ML, Houdebine LM & Bosze Z 1996 Structure of the rabbit κ-casein encoding gene: expression of the cloned gene in the mammary gland of transgenic mice. *Gene* **174** 27–34

Caroli A, Jann O, Budelli E, Bolla P, Jägger S & Erhardt C 2001 Genetic polymorphism of goat κ-casein (*CSN3*) in different breeds and characterization at DNA level. *Animal Genetics* **32** 226–230

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. *Journal of Dairy Science* **87** 2606–2613

Damiani G, Florio S, Budelli E, Bolla P & Caroli A 2000 Single nucleotide polymorphisms (SNPs) within Bov-A2 SINE in the second intron of bovine and buffalo k-casein (CSN3)gene. *Animal Genetics* **31** 277–279

Di Luccia A, Mauriello R, Chianese L, Moio L & Addeo F 1990 κ-Casein polymorphism in caprine milk. *Scienza e Tecnica Lattiero-Casearia* **41** 305–314

Di Gregorio P, Rando A, Pieragostini E & Masina P 1991 DNA polymorphism at the casein loci in sheep. *Animal Genetics* **22** 21–30

Edlund A, Johansson T, Leidvik B & Hansson L 1996 Structure of the human kappa-casein gene. *Gene* **174** 65–69

Feligini M, Cubric Curik V, Parma P, Curik I, Greppi GF & Enne G 2002 Polymorphism of κ-casein in Italian goat breeds: A new ACRS-PCR designed DNA test for discrimination of A and B alleles. *Food Technology and Biotechnology* **40** 293–298

Ginger MR & Grigor MR 1999 Comparative aspects of milk caseins. *Comparative Biochemistry and Physiology B* **124** 133–145

Grosclaude F 1988 [Genetic polymorphism of the main cow milk proteins. Relations with milk's quantity, composition and cheesemaking attitude]. INRA. *Productions Animales* **1** 5–17

Jann O, Prinzenberg EM, Luikart G, Caroli A & Erhardt G 2004 High polymorphism in the κ-casein (*CSN3*) gene from wild and domestic caprine species revealed by DNA sequencing. *Journal of Dairy Research* **71** 188–195

Leveziel H, Metenier L, Guerin G, Cullen P, Provot C, Bertaud M & Mercier JC 1991 Restriction fragment polymorphism of ovine casein gene: close linkage between the α_{s1}-, α_{s2}-, β- and κ-casein loci. *Animal Genetics* **22** 1–10

Lodes A, Krause I, Buchberger J, Aumann J & Klostermeyer H 1996 The influence of genetic variants of milk proteins on the compositional and technological properties of milk. 1. Casein micelle size and the content of non-glycosylated k-casein. *Milchwissenschaft* **51** 368–373

Macheboeuf D, Coulon JB & D'Hour P 1993 Effect of breed, protein genetic variants and feeding on cows' milk coagulation properties. *Journal of Dairy Research* **60** 43–54

Maddox JF, Davies KP, Crawford AM, Hulme DJ, Vaiman D, Cribiu EP, Freking BA, Beh KJ, Cockett NE, Kang N, Riffkin CD, Drinkwater R, Moore SS, Dodds KG, Lumsden JM, Van Stijn TC, Phua SH, Adelson DL, Burkin HR, Broom JE, Buitkamp J, Cambridge L, Cushwa WT, Gerard E, Galloway SM, Harrison B, Hawken RJ, Hiendleder S, Henry HM, Medrano JF, Paterson KA & Schibler L 2001 An enhanced linkage map of the sheep genome comprising more than 1000 loci. *Genome Research* **11** 1275–1289.

Martin P 1993 [Genetic polymorphism of goat milk proteins]. *Lait* **73** 511–532

Mercier JC & Vilotte JL 1993 Structure and function of milk protein genes. *Journal of Dairy Science* **76** 3079–3098

Ng-Kwai-Hang KF 1998 Genetics polymorphism of milk protein: relationships with production traits, milk composition and technological properties. *Canadian Journal of Animal Science* **78** 131–147

Ng-Kwai-Hang KF, Monardes HG & Hayes JF 1990 Association between genetic polymorphism of milk proteins and production traits during three lactations. *Journal of Dairy Science* **73** 3414–3420

Ordás JG, Rando A, Senese C & Masina P 1997 DNA polymorphisms of caseins genes in Spanish dairy sheep. *Small Ruminant Research* **26** 9–13

Parsons BL & Heflich RH 1997 Genotypic selection methods for the direct analysis of point mutations. *Mutation Research* **387** 97–121

Pérez-Rodríguez ML, Martín-Alvarez PJ, Ramos M, García-Muro E, Zaraga I & Amigo L 1998 Genetic polymorphism of bovine milk proteins in Holstein-Friesian and Fleckvieh breeds exploited in Spain. *Milchwissenschaft* **53** 543–546

Prinzenberg EM, Krause I & Erhardt G 1999 SSCP analysis at the bovine CSN3 locus discriminates six alleles corresponding to known protein variants (A, B, C, E, F, G) and three new DNA polymorphisms (H, I, A1). *Animal Biotechnology* **10** 49–62

Rahali V & Ménard JL 1991 [Influence of β-lactoglobulin and κ-casein genetic variants on milk composition and cheesemaking attitude]. *Lait* **71** 275–297

Rajeevan MS, Ranamukhaarachchi DG, Vernon SD & Unger ER 2001 Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. *Methods* **25** 443–451

Russo V, Davoli R, Dall'Olio S & Tedeschi M 1986 [Research on protein polymorphism from goat milk]. *Zootecnica e Nutrizione Animale* **12** 55–62

Sanders Sevall J 2001 Rapid allelic discrimination from real-time DNA amplification. *Methods* **25** 452–455

Soria LA, Iglesias GM, Huguet MJ & Mirande SL 2003 A PCR-RFLP test to detect allelic variants of the bovine kappa-casein gene. *Animal Biotechnology* **14** 1–5

Vegarud GE, Devold TG, Opheim R, Loeding E, Svenning C, Abrahamsen RK, Lien S & Langsrud T 1999 Genetic variants of Norwegian goats milk composition, micellar size and renneting properties. *International Dairy Journal* **9** 367–368

Waugh DF 1971 Formation and structure of micelles. In *Milk Proteins: Chemistry and Molecular Biology*, pp. 4–85 (Ed. HA McKenzie). New York: Academic Press

Wittwer CT, Herrmann MG, Gundry CN & Elenitoba-Johnson KSJ 2001 Real time multiplex PCR assays. *Methods* **25** 430–442

https://doi.org/10.1017/S0022029905000932 Published online by Cambridge University Press