PAPER

Molecular characterisation of κ–*casein* gene in *Girgentana* dairy goat breed and identification of two new alleles

Rosalia Di Gerlando, Lina Tortorici, Maria T. Sardina, Giuseppina Monteleone, Salvatore Mastrangelo, Baldassare Portolano

Dipartimento Scienze Agrarie e Forestali, Università di Palermo, Italy

Abstract

The κ-casein fraction plays an important role in the formation, stabilisation and aggregation on casein micelles and thus affects technological and nutritional properties of milk. In this study, exon 4 of *κ-casein* (*CSN3*) gene was sequenced and analysed in *Girgentana* goat breed. Analyses of the obtained sequences showed the presence of *A*, *B*, *D*, and *G* known alleles and two new genetic variants, named *D'* and *N*. The new *D'* allele differs from *D* in one transition, $G_{284}\rightarrow A_{284}$, which did not cause amino acid change. The new *N* allele differs from *A* in five single nucleotide polymorphisms (SNPs): T_{245}/C_{245}, G_{284}/A_{284}, G_{309}/A_{309}, G_{471}/A_{471} and T_{591}/C_{591}, while it differs from *C* in one transition, *i.e.* $T_{583}\rightarrow C_{583}$. Comparing the amino acid sequences of *N* and *A* alleles, the first two SNPs caused no amino acid change, whereas the other SNPs produced changes (Val_{65}/Ile_{65}, Val_{119}/Ile_{119}, and Ser_{159}/Pro_{159}, respectively). Comparison of *N* allele with *C* revealed the amino acid change $Val_{156}\rightarrow Ala_{156}$. The most frequent allele was *A* (0.480) followed by *B* (0.363), *D* (0.112), and *N* (0.034). The *D'* and *G* alleles were identified only in two animals and in heterozygous conditions with a very low frequency (0.005). The most common genotype was *AB* (39.5%) followed by *AA* (19.5%), *AD* (12.7%), and *BB* (11.7%). Homozygous *D'D'*, *GG*, and *NN* individuals were not found. Further analysis will be performed in order to establish associations among genotypes and quantitative and qualitative milk traits.

Introduction

In the milk of ruminants, more than 95% of proteins are synthesised by six structural genes, four caseins [αs_1-, β-, αs_2- and *κ-casein* (*CSN3*)] and two whey proteins (*α-lactalbumin* and *β-lactoglobulin*). Polymorphisms of the four casein genes have been the focus of considerable research effort because of their potential effects on milk quality. The κ-casein fraction plays an important role in the formation, stabilisation and aggregation of the casein micelles and thus affects the technological (Mariani *et al*., 1976; Aleandri *et al.*, 1990; Lodes *et al.,* 1996; Falaki *et al*., 1997) as well as nutritional properties of milk (Mercier *et al*., 1973, 1976; Malkoski *et al*., 2001).

The goat *CSN3* gene comprises five exons (Coll *et al.,* 1993, 1995) with the mRNA coding region for mature protein (171 amino acids) spanning from exon 3 (9 amino acids) to exon 4 (162 amino acids) (Yahyaoui *et al*., 2003). The *CSN3* gene is considered to be monomorphic in sheep (Moioli *et al*., 1998) whereas several studies on goat *CSN3* showed that this gene is highly polymorphic (Caroli *et al.,* 2001; Yahyaoui *et al.,* 2001; Angiolillo *et al*., 2002; Chessa *et al*., 2003; Yahyaoui *et al*., 2003; Jann *et al*., 2004; Reale *et al*., 2005; Prinzenberg *et al*., 2005; Gupta *et al*., 2009; Kiplagat *et al*., 2010). According to the last nomenclature proposed by Prinzenberg *et al*. (2005), a total of 16 DNA variants have been identified in the domestic goat, of which 13 are protein variants (named in alphabetical order from *A* to *M*) and 3 are silent mutations (*B'*, *B''* and *C'*) involving a total of 15 polymorphic sites. Recently, Gupta *et al*. (2009) in Jakhrana goat breed and Kiplagat *et al*. (2010) in indigenous Eastern African goat population reported the presence of new genetic polymorphisms at *CSN3* gene. These studies reported conflicting results for allele nomenclature because according to Prinzenberg *et al*. (2005) only missense mutations associated with amino acid changes should be indicated with new allele names (*i.e.* new letter), while silent mutations should be named with the same letter as the related protein-associated allele followed by prime symbol. The *Girgentana* goat is an ancient Sicilian goat breed reared in Southern Italy for its good dairy production. Average milk production was 224±66 L in the first lactation, and 320±109 L for later lactations (AIA, 2013). Due to sanitary policies, population size of *Girgentana* goat breed decreased of almost 90% in 20 years. In 1983, the population consisted of 30,000 individuals but, nowadays, only 374 heads are enrolled in the Herd Book (ASSONAPA, 2013).

Corresponding author: Prof. Baldassare Portolano, Dipartimento Scienze Agrarie e Forestali, Università di Palermo, viale delle Scienze 4, 90128 Palermo, Italy.
Tel. +39.091.23896069 - Fax: +39.091.23860814.
E-mail: baldassare.portolano@unipa.it

Key words: *CSN3* gene, Polymorphisms, New alleles, *Girgentana* goat breed.

Acknowledgments: this work was supported by Rural Development Program 2007-2013, Sicily Region, Italy, Measure 1.2.4, Grant N. CUP:G66D11000039999.

Received for publication: 20 May 2014.
Accepted for publication: 29 November 2014.

Italian Journal of Animal Science 2015; 14:3464
doi:10.4081/ijas.2015.3464

Over the last years this breed has become almost extinct, in part as a consequence of the marked decrease in fresh goat milk consumption. The aim of this work was to investigate the genetic polymorphisms of *CSN3* gene in the *Girgentana* dairy goat breed in order to assess genotypes distribution and to use this information in future conservation programmes for this breed considering that genotype could influence milk properties.

Materials and methods

Sampling and DNA extraction

A total of 205 individuals, all females, of *Girgentana* goat breed were randomly chosen. They belonged to 15 different herds located in different areas of Sicily. Samples were collected from 10 to 20 unrelated individuals per herd. About 10 mL of blood was collected from jugular vein using vacutainer tubes containing ethylenediamine tetraacetic acid (EDTA) as anticoagulant. Genomic DNA was extracted from buffy coats of nucleated cells using a salting out method (Miller *et al.,* 1988). After checking the quantity and quality of the DNA using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), samples were diluted to a final concentration of 50 ng/μL in ultrapure water and stored at 4°C until use.

DNA amplification and purification

A 552 bp fragment of *Girgentana* goat *κ-casein* exon 4 (GenBank Acc. No. X60763 mRNA goat *CSN3*) was amplified by polymerase chain reaction (PCR) using the following primers: forward -AGAAATAATACCATTCTGCAT, and reverse - TCTTTGATGTCTCCTTAGAG. The PCR reaction was performed in a 25 µL of final volume containing 1 µM of each primer, 1 mM of dNTP Mix, 1 U of Taq DNA polymerase (Fermentas, Hanover, MD, USA), 1X PCR buffer with KCl, 1.25 mM $MgCl_2$, and approximately 100-150 ng of genomic DNA. Thermal cycling conditions were 94°C for 90 sec for initial denaturation, 30 cycles of 45 sec each at 94°C, 50°C and 72°C, with a final extension at 72°C for 5 min. The PCR products were checked by electrophoresis on 2% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, CA, USA).

DNA sequencing reaction

All the collected samples were amplified and the PCR products purified in order to sequence and determine the complete nucleotide sequences. Polymerase chain reaction products were purified using 10 U of Exonuclease I and 1 U of Shrimp Alkaline Phosphatase (Fermentas). DNA sequencing reaction was carried out using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA) with 5 µM of the same primers used in the PCR reaction. Cycle sequencing reaction was performed according to manufacturer's instruction following Ethanol/EDTA/Sodium Acetate precipitation. Sequencing analyses were performed in an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems).

Sequence data analysis

Nucleotide sequences obtained were checked using Sequencing Analysis Software v5.3.1 (Applied Biosystems) and subsequently analysed with SeqScape v2.5 Software (Applied Biosystems). Polymorphic sites were confirmed by visual examination of the electropherograms. Multiple alignments of the sequences were performed using ClustalW software (Thompson *et al*., 1994). The translation of DNA sequences to amino acid sequences was performed using ExPASy-Traslate tool. The same software was used to calculate the isoelectric point (IP) of the new genetic variants found in *Girgentana* goat breed. The exact P value associated with the null hypothesis of Hardy-Weinberg equilibrium (HWE) was estimated using GENEPOP version 4.0.11 (Rousset, 2008). The programme performed a probability test using a Markov Chain

Table 1 κ-casein gene variants according to Prinzenberg *et al.* (2005).

CSN3 variant	IEF pattern	GenBank Acc. no.	Nucleotide position														Reference
			170	245	247	274	284	290	298	309	384	385	471	509	550	583	591	
A	A^{IEF}	X60763	C	T	A	A	G	C	A	G	G	A	G	A	T	C	T	Coll *et al.* (1993)
B	A^{IEF}	AF485340											A					Yahyaoui *et al.* (2001)
		AF434988											A					Jann *et al.* (2004)
B'	A^{IEF}	AY166706	T										A					Jann *et al.* (2004)
B'	A^{IEF}	AY166707						T					A					Jann *et al.* (2004)
C	A^{IEF}	AY350425		C			A			A			A			T(Val)	C	Prinzenberg *et al.* (2005)
C'	A^{IEF}	AF485341		C			A			A			A	G		T	C	Yahyaoui *et al.* (2001)
D	B^{IEF}	AY027868		C	G					A			A				C	Caroli *et al.* (2001)
		AY090465		C	G		G(Leu)			A			A				C	Yahyaoui *et al.* (2001)
D'	$B^{IEF\circ}$	JX889422		C	G		A(Leu)			A			A				C	This paper
E	B^{IEF}	AF486523										G	A					Angiolillo *et al.* (2002)
F	A^{IEF}	AY090466		C									A				C	Yahyaoui *et al.* (2003)
G	A^{IEF}	AY090467		C						A			A				C	Yahyaoui *et al.* (2003)
H	A^{IEF}	AF521022				G							A					Jann *et al.* (2004)
I	A^{IEF}	AY166710								A			A					Jann *et al.* (2004)
J	A^{IEF}	AY166711							G				A					Jann *et al.* (2004)
K	B^{IEF}	AY166709			G								A					Jann *et al.* (2004)
L	A^{IEF}	AY166708								A			A				C	Jann *et al.* (2004)
M	B^{IEF}	AY428577		C							A		A		C		C	Prinzenberg *et al.* (2005)
N	$A^{IEF\circ}$	JX889424		C			A			A			A			C(Ala)	C	This paper

CSN3, *κ-casein* gene; IEF, isoelectrofocusing. °The IEF pattern of D' and N alleles was not experimentally tested but it was estimated using Expasy pI tool. GenBank accession numbers, nucleotide position compared with X60673, and published reference are indicated. Amino acids are indicated within brackets. Reproduced with permission of the American Dairy Science Association via Elsevier: Journal of Dairy Science, Vol. 88, E.-M. Prinzenberg, K. Gutscher, S. Chessa, A. Caroli, G. Erhardt, Caprine κ-casein (CSN3) polymorphism: new developments in molecular knowledge, page 1492, Copyright (2005).

method (1000 dememorisation steps, 100 batches, and 1000 iterations per batch). Moreover, GENEPOP was used to calculate genotype and allele frequencies and fixation index F_{is} (Weir and Cockerham, 1984). Expected (*He*) and observed (*Ho*) heterozygosity were calculated using the POPGENE software version 1.31 (Yeh *et al.*, 1999).

Results and discussion

Identified alleles in *Girgentana* goat breed

Sequencing analysis and alignment of the obtained sequences of *CSN3* exon 4 showed the presence in *Girgentana* goat breed of *A*, *B*, *D*, and *G* known alleles and two new genetic variants (GenBank Acc. No. JX889419-JX889424). All single nucleotide polymorphisms (SNPs) described by Prinzenberg *et al.* (2005), including the two new polymorphic sites detected in our samples, are showed in Table 1. Considering the conflicting results reported by Gupta *et al.* (2009) and Kiplagat *et al.* (2010), we named *D'* and *N* the two new alleles identified in *Girgentana* goat breed according to Prinzenberg *et al*. (2005). The new *CSN3 D'* (GenBank Acc. No. JX889422) allele differing from *CSN3 D* (GenBank Acc. No. AY027868) in one transition $G_{284}\rightarrow A_{284}$, which did not cause amino acid change (Leu_{56}/Leu_{56}). The new *CSN3 N* (GenBank Acc. No. JX889424) allele differing from *CSN3 A* (GenBank Acc. No. X60763) allele in five SNPs (T_{245}/C_{245}, G_{284}/A_{284}, G_{309}/A_{309}, G_{471}/A_{471} and T_{591}/C_{591}), while differing from *C* (GenBank Acc. No. AY350425) allele in one transition ($T_{583}\rightarrow C_{583}$). Comparing the amino acid sequences of *CSN3 N* and *A* alleles, the first two SNPs (T_{245}/C_{245} and G_{284}/A_{284}) caused no amino acidic change, whereas the other SNPs produced changes: Val_{65}/Ile_{65}, Val_{119}/Ile_{119}, and Ser_{159}/Pro_{159}, respectively. Comparison of *CSN3 N* allele with *CSN3 C* allele revealed the amino acid change $Val_{156}\rightarrow Ala_{156}$.

Allele frequencies and genetic variability

Table 2 shows genotype and allele frequencies at *CSN3* locus in *Girgentana* goat breed. The most frequent allele was *A* (0.480) followed by *B* (0.363), *D* (0.112), and *N* (0.034). The *D'* and *G* alleles were identified only in two animals and in heterozygous conditions with a very low frequency (0.005). These results are not in agreement with those reported for *Girgentana* goat breed by Gigli *et al*. (2008), and by other authors for Italian (Sacchi *et al.*, 2005), European and African (Prinzenberg *et al*., 2005) goat breeds, where the most frequent allele was *B*. Prinzenberg *et al*. (2005), proposed to differentiate the nomenclature at protein level from the one used at DNA level introducing two codes (A^{IEF} and B^{IEF}) corresponding to two IPs (IP=5.53 and 5.78, respectively) identified using isoelectrofocusing (IEF) method (Table 1). According to this nomenclature, among the *CSN3* alleles found in our study, only the *D* and *D'* alleles are included in B^{IEF} group, whereas *A*, *B*, *G* and *N* alleles belong to A^{IEF} group (Table 1), which represents the less favourable variants group in terms of milk composition and technological properties (Chiatti *et al*., 2007).

Table 2. Genotype and allele frequencies at *κ-casein* locus in *Girgentana* goat breed.

Genotype (n)	Frequency	Allele	Frequency
AA (40)	0.195	A	0.480
AB (81)	0.395	B	0.363
BB (24)	0.117	D	0.112
AD (26)	0.127	D'	0.005
BD (16)	0.078	G	0.005
DD (2)	0.010	N	0.034
D'G (2)	0.010		
AN (10)	0.049		
BN (4)	0.020		

n, number of individuals.

Nine genotypic classes were found in our *Girgentana* goat samples. The most common genotype was *AB* (39.5%) followed by *AA* (19.5%), *AD* (12.7%), and *BB* (11.7%). The other genotypes showed a frequency of less than 10% (Table 2). In this study, we found no homozygous *D'D'*, *GG*, and *NN* subjects. Caravaca *et al.* (2009), in a study on the effect of *CSN3* genotypes on goat milk composition, showed that *AB* and *BB* genotypes were significantly associated with higher levels of total casein and protein content compared with the *AA* genotype, thus underlining the importance of taking into account the *CSN3* genotype when performing selection for milk composition in dairy goats.

Observed and expected heterozygosity, fixation index F_{is} and P value associated with the null hypothesis of HWE were estimated. Significant departure from HWE was observed for *Girgentana* goat breed at *CSN3* locus ($P<0.05$), probably due to heterozygote excess (*Ho*=0.6766 *vs* *He*=0.6243). This hypothesis could be confirmed considering *Ho* heterozygosity and F_{is} (-0.0855) values.

Conclusions

Two new genetic variants have been identified and characterised in *Girgentana* goat breed. Currently, phenotypic data are not available for this goat breed; hence, further studies could establish the possible association and the effects of polymorphisms on quantitative and qualitative characteristics of milk. Moreover, it could be useful to take into account *CSN3* gene to use lines of goats producing different types of milk for specific cheese-making technologies or nutritional human needs.

References

AIA, 2013. Italy: milk recording activity. Official statistics. Italian Breeders Association Publ., Rome, Italy.

Aleandri, R., Buttazzoni, L.G., Schneider, J.C., Caroli, A., Davoli, R., 1990. The effects of milk protein polymorphisms on milk components and cheese-producing ability. J. Dairy Sci. 73:241-255.

Angiolillo, A., Yahyaoui, M.H., Sanchez, A., Pilla, F., Folch, J.M., 2002. Short communication: characterization of a new genetic variant in the caprine κ-casein gene. J. Dairy Sci. 85:2679-2680.

ASSONAPA, 2013. Available from: http://www.assonapa.it/norme_ecc/Consistenze_Caprini.htm

Caravaca, F., Carrizosa, J., Urrutia, B., Baena, F., Jordana, J., Amills, M., Badaoui, B., Sanchez, A., 2009. Short communication: effect of α_{S1}-casein (CSN1S1) and κ-casein (CSN3) genotypes on milk composition in Murciano-Granadina goats. J.

Dairy Sci. 92:2960-2964.
Caroli, A., Jann, O., Budelli, E., Bolla, P., Jäger, S., Erhardt, G., 2001. Genetic polymorphism of goat κ-casein (CSN3) in different breeds and characterization at DNA level. Anim. Genet. 32:226-230.
Chessa, S., Budelli, E., Gutscher, K., Caroli, A., Erhardt, G., 2003. Short communication: simultaneous identification of five kappa-casein (CSN3) alleles in domesticated goat by polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP). J. Dairy Sci. 86:3726-3729.
Chiatti, F., Chessa, S., Bolla, P., Cigalino, G., Caroli, A., Pagnacco, G., 2007. Effect of the κ-casein polymorphism on milk composition in the Orobica goat. J. Dairy Sci. 90:1962-1966.
Coll, A., Folch, J.M., Sanchez, A., 1993. Nucleotide sequence of the goat kappa-casein cDNA. J. Anim. Sci. 71:2833.
Coll, A., Folch, J.M., Sanchez, A., 1995. A structural features of the 5' flanking region of the caprine kappa casein gene. J. Dairy Sci. 78:973-977.
Falaki, M., Prandi, A., Corradini, C., Sneyers, M., Gengker, N., Massart, S., Fazzini, U., Burny, A., Portetelle, D., Renaville, R., 1997. Relationship of growth hormone geneand milk protein polymorphisms to milk production traits in Simmental cattle. J. Dairy Res. 64:47-56.
Gigli, I., Maizon, D.O., Riggio, V., Sardina, M.T., Portolano, B., 2008. Casein haplotype variability in Sicilian dairy goat breeds. J. Dairy Sci. 91:3687-3692.
Gupta, S.C., Kumar, D., Pandey, A., Malik, G., Gupta, N., 2009. New k-casein alleles in Jakhrana goat affecting milk processing properties. Food Biotechnol. 23:1-14.
Jann, O.C., Prinzenberg, E.M., Luikart, G., Caroli, A., Erhardt, G., 2004. High polymorphism in the kappa-casein (CSN3) gene from wild and domestic caprine species revealed by DNA sequencing. J. Dairy Res. 71:188-195.
Kiplagat, S.K., Agaba, M., Kosgey, I.S., Okeyo, M., Indetie, D., Hanotte, O., Limo, M.K., 2010. Genetic polymorphism of kappa-casein gene in indigenous Eastern Africa goat populations. Int. J. Genet. Mol. Biol. 2:1-5.
Lodes, A., Buchberger, J., Aumann, J., Klostermeyer, H., 1996. The influence of genetic variants of milk proteins on the compositional and technological properties of milk. Casein micelle size and the content of non-glycosylated kappa-casein. Milchwissenschaft 51:368-373.
Malkoski, M., Dashper, S.G., O'Brien-Simpson, N.M., Talbo, G.H., Macris, M., Cross, K.J., Reynolds, E.C., 2001. Kappacin, a novel antibacterial peptide from bovine milk. Antimicrob. Agents Ch. 45:2309-2315.
Mariani, P., Losi, G., Russo, V., Castagnetti, G.B., Grazia, L., Morini, D., Fossa, E., 1976. Prove di caseificazione con latte caratterizzato dalle varianti A e B della nella produzione del formaggio Parmigiano-Reggiano. Sci. Tecn. Latt. Cas. 27:208-227.
Mercier, J.C., Brignon, G., Ribadeau-Dumas, B., 1973. Primary structure of bovine kappa B casein. Complete sequence. Eur. J. Biochem. 35:222-235.
Mercier, J.C., Chobert, M., Addeo, F., 1976. Comparative analysis of the amino acid sequences of caseinomacropeptide from seven species. FEBS Lett. 72:208-214.
Miller, S.A., Dykes, D.D., Polesky, H.F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 16:1215.
Moioli, B., D'Andrea, S., Pilla, F., 1998. Candidate gens affecting sheep and goat milk quality. Small Ruminant Res. 68:179-192.
Prinzenberg, E.M., Gutscher, K., Chessa, S., Caroli, A., Erhardt, G., 2005. Caprine κ-casein (CSN3) polymorphism: new developments in molecular knowledge. J. Dairy Sci. 88:1490-1498.
Reale, S., Yahyaoui, M.H., Folch, J.M., Sanchez, A., Pilla, F., Angiolillo, A., 2005. Genetic polymorphism of the κ-casein (CSN3) gene in goats reared in Southern Italy. Ital. J. Anim. Sci. 4:97-101.
Rousset, F., 2008. Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. Mol. Ecol. Resour. 8:103-106.
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. J. Dairy Sci. 88:1561-1568.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.
Yahyaoui, M.H., Angiolillo, A., Pilla, F., Sanchez, A., Folch, J.M., 2003. Characterization and genotyping of the caprine kappa casein variants. J. Dairy Sci. 86:2715-2720.
Yahyaoui, M.H., Coll, A., Sanchez, A., Folch, J.M., 2001. Genetic polymorphism of the caprine kappa casein gene. J. Dairy Res. 68:209-216.
Yeh, F.C., Yang, R., Boyle, T., 1999. Popgene v. 1. 31. Microsoft Windows-based freeware for population genetic analysis. University of Alberta and Centre for International Forestry Research Publ., Alberta, Canada.