Identification of the CSN1S1 allele in Indian goats by the PCR-RFLP method

A. Kumar, P. K. Rout‡, A. Mandal and R. Roy

Genetics and Breeding Division, CIRG, Makhdoom, Farah, Mathura 281122, Uttar Pradesh, India

(Received 16 January 2007; Accepted 14 June 2007)

The allelic distributions of the CSN1S1 (αs1-casein) in the Indian goats are quite different from European goat breeds. Majority of Indian goat breeds and non-descript goats carry A, B, E and F alleles at αs1-casein locus, as found by analysing both DNA and protein levels. However, A and B alleles, known to be associated with better casein yield, were observed in the highest proportion in all the Indian goat breeds. Gene frequency and breed heterozygosity were computed for the CSN1S1 gene. The gene frequency of allele A in Indian goats varies from 0.68 to 1.00 and allele B varies from 0.098 to 0.23. Allele F was observed in Beetal, Marwari, Chegu and non-descript goats of MP (Local MP) in less than 1% of population. The expected heterozygosity (He) varied from 0.141 to 0.506 over the population. The Beetal breed showed the highest gene diversity (0.506) followed by Jamunapari (0.395), Chegu (0.383) and Jakhroma (0.381) breeds. Therefore, the variability at CSN1S1 locus can be utilised for conservation as well as for genetic improvement of Indian goat breeds for increasing both the quality and quantity of milk production.

Keywords: casein, genetic diversity, goats, polymorphism, RFLP

Introduction

Casein, the main protein fraction of ruminant milk, is one of the most valuable components due to its nutritional value and processing properties. The casein fraction is encoded by four tightly linked genes and is organised as a cluster in a 250-kb genomic DNA segment in the following order: αs2-casein (CSN1S2), αs1-casein (CSN1S1), β-casein (CSN2) and κ-casein (CSN3) (Grosclaude et al., 1987; Ferretti et al., 1990; Threadgill and Womack, 1990). They have been mapped on chromosome 6 in cattle and goats (Hayes et al., 1993; Popescu et al., 1996). The existence of extensive polymorphism at αs1-casein provided unusual quantitative and qualitative differences in casein synthesis (Boulanger et al., 1984; Grosclaude et al., 1987; Ramunno et al., 1991) as well as its importance to the application of milk protein research to dairy industry (Bevilacqua et al., 2002; Feligini et al., 2005). The αs1-casein, which has known 16 co-dominant alleles, is associated with different rates of protein synthesis. It has been established that the A, B and C alleles at CSN1S1 locus associated with the production of a high level of protein in milk, the E allele associated with a medium-level protein and F and G alleles with a low-level protein (Grosclaude et al., 1987 and 1994). The distribution of different alleles at CSN1S1 locus has been investigated in European countries at the genomic as well as protein levels (Grosclaude et al., 1994, Ramunno et al., 1991 and 2000; Bevilacqua et al., 2002; Sacchi et al., 2005). However, casein variability has not been well characterised in Indian goats except for the reports of Prakash et al. (2002) and Rout et al. (2004a). Moreover, milk protein gene diversity in cattle showed a strong relationship with human lactase gene and has been used to analyse geographical and genetic diversity in European cattle (Beja-Pereira et al., 2003). India has a high goat population, which is distributed over all the regions of the country, and produce significant amount of goat milk. The characterisation of CSN1S1 casein variability is important due to its relationship with cheese production and milk-processing properties. Therefore it is important to evaluate the variability of αs1-casein gene in different Indian goat breeds and to determine the effect of αs1-casein genotyping for genetic improvement as well as for a conservation programme. Therefore the present study has been designed to characterise αs1-casein (CSN1S1 locus) in Indian goats by both genomics and protein level.

‡ E-mail: rout_ctc@hotmail.com, pramod@cirg.res.in
Material and methods

Sample collection and DNA isolation
A total of 347 unrelated blood samples were collected from goats in their natural habitats belonging to the four major geographical agro-climatic zones of India, including at least one breed from each major geographical region. An effort was made to collect samples from unrelated individuals based on the information provided by farmers. DNA was isolated from the samples using the standard protocol published elsewhere (Thangaraj et al., 2002).

We analysed 13 different genetic groups from various agro-climatic zones for CSN1S1 gene variability by DNA analysis (Table 1). As we need to correlate structural variation with functional variation in different goat breeds, genotyping was carried out to analyse milk protein variability at CSN1S1 locus. Therefore, we collected milk samples from all major milk-producing goat breeds namely Jamunapari, Barbari, Marwari, Sirohi, Jakhrana and Beetal belonging to the arid and semi-arid climatic zones, and from two non-descript goats (Local Madhya Pradesh (MP) and Local Uttar Pradesh (UP)) available in the same agro-climatic area (Table 2). Milk protein analysis was carried out to provide supporting evidence on variation at the DNA level.

Milk samples were collected from 1058 goats of Jamunapari, Barbari, Marwari, Sirohi, Jakhrana, Beetal, Local UP and Local MP goats from the natural habitat of each breed and transported to the laboratory and stored at −20°C for further analysis.

Genetic stocks
Jamunapari goats are found in isolated pockets and sampling was carried out in their natural habitat in the Chakarnagar area of Etawah district of UP. The breed is known as the best Indian dairy goat (Rout et al., 2004b) and the breed risk status was classified as endangered (http://clid.fao.org) (Rout et al., 2000). Similarly, Jakhrana is known for high milk yield from the Alwar area of Rajasthan. Goats are distributed in few villages in their home tract and the number is also decreasing in the natural habitat. Barbari is a medium-sized breed of the semi-arid zone and known for its adaptability over a wide range of agro-climatic situations. Beetal is one of the largest breeds of goat and reared mainly for milk but is equally important for meat. This breed is found in Punjab along the Indo-Pakistan border and the number is decreasing in the region. Sirohi is a medium- to large-sized breed and is best known for meat, milk production and resistance to a number of diseases. The natural habitat of the breed is in the wider area of Ajmer, Bhilwara, Tonk and Jaipur in Rajasthan. Marwari is a medium-sized breed with a compact body and strong legs and is known for its hardness and adaptability to extreme temperatures. Marwari is also known for meat, milk and coarse fibre production. The Black Bengal is the typical dwarf breed of eastern India and known for its high prolificacy and meat quality. Osmanabadi goats are a medium-sized breed with a comparatively long body and long legs and are found in the Ahmednagar and Solapur area of Maharashtra. Gaddi goats are adapted to a hilly environment, found in the hills of Himachal Pradesh and are known for their draftability. Chegu goats are known for Pashmina production and are found at high altitudes of Lahaul and Spliti Valley of Himachal Pradesh, Uttarkashi, Chamoli and Pithoragarh district of Uttarakhand. Surti goats are found predominantly in the Bhavnagar area of Gujarat and is known for its milk production. ‘Local’ goats are non-descript goats found in the area adjacent to Central Institute for Research on Goats (CIRG), Makhdoom and Hathras area of UP, designated as Local UP, and from adjacent region of Gwalior and Morena area of MP, designated as Local MP. These goats are mainly dual-purpose goats and reared by local farmers.

Genotyping
The variation at the DNA level was analysed in 13 genetic groups (Table 1). Milk protein variation was measured in seven populations to provide supporting evidence for genomic variability (see Table 2). DNA samples were analysed with allele-specific polymerase chain reaction (AS-PCR) and the amplified product was digested with XmnI for the 13 genetic groups. PCR was carried out in a 50 µl reaction mixture containing 100 ng genomic DNA, 10 pmol of each primer (forward: F 5’ TACTTGAAGCTCA GAGGCAG-3’, reverse: 5’ GGATGATAGCCTGTGTG-3’), 1.25 U of TaqDNA polymerase, 50 mmol/l KCl, 10 mmol/l Tris–HCl (pH 9.0), 0.1% Triton X-100, 3 mmol/l MgCl2, dNTPs each at 400 µmol/l, 0.04% BSA (Ramunno et al., 2000). The amplification protocol was used as follows: an initial cycle of 97°C for 2 min, 60°C for 45 s and 72°C for 2 min 30 s; then 30 cycles of 94°C for 45 s, 60°C for 45 s and 72°C for 2 min 30 s and a final extension step 72°C for 10 min. Restriction analysis was carried out using XmnI enzyme. For this, 20 µl of each PCR product was digested with 10 U of XmnI endonuclease for overnight at 37°C and digested products were analysed in 4% agarose gel stained with ethidium bromide and analysed in a gel documentation system (Alpha Innotech Corporation, San Leandro, CA, USA).

SDS-PAGE was carried out in skimmed milk samples and was analysed by means of SDS (Grosclaude and Martin, 1997) and urea (Medrano and Sharrow, 1989) PAGE and alkaline pH. Gels were stained with Coomassie Brilliant Blue. Milk protein variants were determined by the molecular weight in gel documentation system (Alpha Innotech Corporation).

Statistical analysis
Genepop (Raymond and Rousset, 1995) software was used to estimate allelic frequencies, expected heterozygosity, effective number of alleles and to verify Hardy–Weinberg equilibrium. The genotypes were observed by counting the patterns in the gel documentation system. Variance and
95% confidence interval were calculated according to the formula \( \text{Var}(\hat{p}_u) = \frac{1}{N} \left( \hat{p}_u (1 - \hat{p}_u) - \frac{\hat{p}_u^2}{N} \right) \) suggested by Weir (1996) \((\hat{p}_u = \text{genotypic frequency}; \hat{p}_a = \text{allelic frequency}).

### Results

DNA samples were analysed for the presence of different \(\alpha_s\)-casein allele by single-AS-PCR. The region of goat \(\alpha_s\)-casein gene between nucleotide 208 and 420 spanning part of eighth intron, ninth exon and part of the ninth intron was amplified and digested with \(XmnI\). Observed genotypic and allelic frequency at the \(\alpha_s\)-casein locus of different Indian goat breeds are presented in Table 1.

Genotyping of 347 individuals belonging to 13 different genetic groups or breeds of Indian goats was carried out with the PCR-RFLP method. The PCR amplified product was observed as 223 bp (Figure 1). Four different variant groups of (150 + 63) bp, (161 + 63) bp, (212 + 150 + 63) bp and (223 + 150 + 63) bp were obtained (Figure 2) after restriction digestion with \(XmnI\). Genotyping at the DNA level showed that the AA genotype had the highest frequency in Indian goats. Comparing results from both DNA and milk samples of individuals revealed the (150 + 63 bp) haplotype in higher proportion in all the breeds except in Local UP goats. This haplotype is associated with the presence of \(\alpha_s\)-casein allele \(\alpha_{S1}^{\alpha_s}\) and \(\alpha_{S1}^{\alpha_s}\) alleles (null allele) at \(\alpha_s\)-casein locus. The percentage of (161 + 63 bp) haplotype varies from 5% to 22.9% in the Indian goat population, indicating a significant proportion of B and E alleles in the population. The electrophoretic pattern in SDS-PAGE did not show the presence of the E variant in the goat breeds. AF genotypes (223 + 150 + 63 haplotype) were observed in Marwari, Beetal, Chegu and Local MP goats indicating the presence of F and A alleles in the population. The F variant was also observed in SDS-PAGE, which was confirmed by the presence of the F allele in Marwari, Beetal, Chegu and Local MP goats. The highest percentage of (212 + 50 + 163bp) haplotype was observed in Beetal, Osmanabadi and Sirohi goats, indicating the presence of the A and D allele in the population and the D allele was not observed in SDS-PAGE. The presence of 11 bp insertion (223 or 224 bp) was also observed in the Indian goats. The absence of 11 bp insertion (212 or 213 bp) was also observed in the Indian goats at the ninth exon. The Indian goats showed the presence of A, B, E and F alleles at \(\alpha_s\)-casein locus by both DNA and protein analysis.

Milk protein analysis of SDS-PAGE pattern revealed a similar type of variability at \(\alpha_s\)-casein locus in the Indian goats and confirmed the findings as obtained from DNA analysis. The \(\alpha_s\)-casein A allele was observed in the majority of goats and their frequency in Jamunapari, Barbari, Marwari, Sirohi, Jakhrama, Beetal, Local UP and Local MP was 0.71, 0.77, 0.56, 0.76, 0.67, 0.72, 0.58 and 0.52, respectively. The B variant was observed as heterozygous...
AB in the Indian goat breeds. The F allele was distributed as heterozygous AF in different goat breeds. (Table 1)

Gene frequencies and their 95% confidence interval were presented in all the studied populations (13 genetic groups; Table 1). Hardy–Weinberg equilibrium (HWE) was tested by $\chi^2$-tests in popgene software and there was no significant departure from Hardy–Weinberg equilibrium ($\chi^2 > 5.246$) observed in the analysed population. The expected heterozygosity (He) varies from 0.141 to 0.506 over the population. The Beetal breed showed highest gene diversity (0.506) followed by Jamunapari (0.395), Chegu (0.383) and Jakhrana (0.381; Table 1). Barbari, Surti and Local MP goats showed lower heterozygosity as compared with other breeds. Similarly, the effective number of alleles varied from 1.00 to 2.02 over the genetic groups. The effective number of allele was highest in Beetal (2.02) followed by Jamunapari, Chegu and Jakhrana (Table 1). No variability was observed at CSN1S1 locus in Gaddi and Local UP samples. Beetal, Jamunapari and Jakhrana breeds showed higher gene diversity and effective number of allele at CSN1S1 locus and are considered as threatened breeds in their natural habitat.

Discussion

CSN1S1 is characterised by 19 exons ranging in size from 24 (exons 5, 6, 7, 8, 10, 13, 16) to 17.5 kb (Jansa Perez et al., 1994). The goat CSN1S1 locus has been characterised by at least 13 alleles, which have been associated with different levels of protein synthesis (Grosclaude et al., 1987; Martin, 1993). Grosclaude et al. (1987) reported that the amount of total casein in caprine milk was positively correlated with the presence of $\alpha_s$1-casein allele and was highest in case of A, B and C alleles. (Boulanger et al., 1984; Grosclaude et al., 1994; Roncada et al., 2002).

Molecular analysis at the DNA level showed A, B, E and F alleles at $\alpha_s$1-casein locus, and A and B alleles were observed in highest proportion in the Indian goat breeds. SDS-PAGE analysis also indicated that the Indian goat breeds are carrying the A allele in higher frequency. Local UP, MP, Gaddi, Black Bengal, Surti, Sirohi and Barbari breeds showed a very high frequency allele A (Table 1). The Beetal breed showed the lowest allele A frequency at this locus, but taking into consideration the 95% confidence interval the frequency varied from 0.513 to 0.837, which is not different from other breeds. However, the sample size of Beetal, Surti, Sirohi and Local MP breeds was less than 25. There is about a 95% chance that the interval included population frequency provided the sample is reasonably large ($n > 30$) (Weir, 1996; Lewis et al., 2004). The frequency of F and E alleles was low in the Indian goats. Allele D was observed in heterozygous forms in the Indian goat breeds by DNA analysis and not observed in SDS-PAGE, which needs further characterisation. Molecular analysis showed the presence of A, B, E and F alleles in the Indian goats and the same was confirmed by SDS-PAGE analysis. SDS-PAGE revealed a lower number allele as it was not possible to determine all variants at the protein level as different variants co-migrate with each other (Grosclaude et al., 1994).

Indian goats are better producers of milk as well as protein in comparison with goat breeds of the other regions. The Italian goat breeds, Garganica and Maltase, exhibited $\alpha_s$1-casein $^A$ frequency as 0.61 and 0.33, respectively, and the Spanish goats, Palmera and Canaria breeds, showed $\alpha_s$1-casein $^A$ frequency as 0.68 and 0.28, respectively (Grosclaude et al., 1987; Ramunno et al., 1991; Jordana et al., 1996). The Spanish goats, Maurciano-Granadina,
Malguena, Payoya and Majorera, showed very low $\alpha_S1$-cn$^A$ frequency (0.05 to 0.28) and higher frequency of $E$ and $B$ alleles (Grosclaude et al., 1987). Alpine and Saanen goats from France showed $\alpha_S1$-cn$^T$ and $\alpha_S1$-cn$^F$ allele frequencies as 0.34 and 0.41, respectively, and $\alpha_S1$-cn$^G$ frequency as 0.14 and 0.07, respectively (Grosclaude et al., 1987 and 1994). Togenburg, Appenzeller and Verzasca breeds of Switzerland had $\alpha_S1$-cn$^G$ frequency as 0.01 for all the breeds and $\alpha_S1$-cn$^F$ frequency was 0.69, 0.44 and 0.62, respectively (Grosclaude et al., 1987 and 1994). The frequency of $\alpha_S1$-cn$^T$ locus was lower in Spanish breeds (0.08, 0.04, 0.0 and 0.0 for Murciano-Granadina, Malguena, Payoya and Canaria, respectively) while the $E$ allele was predominant in Murciano-Granadina (0.59), Malguena (0.65) and Payaya (0.76) breeds (Jordana et al., 1996).

Although Alpine and Saanen produce large amounts of milk, protein content is less as they carry a defective allele in the genome (Grosclaude et al., 1987). The Indian breeds milk, protein content is less as they carry a defective allele.

The analysis indicated there must be conservation of these alleles at DNA level therefore need further characterisation.

Jamunapari, Beetal and Jakhraana breeds exhibited higher gene diversity and effective number of alleles and are presently considered as threatened breeds in their respective home tract. The number of animals is decreasing due to the change in agricultural practices and for other reasons. The analysis indicated there must be conservation of these breeds with respect to milk protein variability. Milk protein diversity showed Beetal as a unique breed indicating the need to conserve the germplasm as the breed is facing extinction in its home tract for several reasons. Moreover, milk protein gene diversity has been analysed in order to explain the cultural evolution of the lactase gene between humans and bovines and also for use in establishing geographical diversity and conservation decisions (Beja-Pereira et al., 2003). Indian goat breeds are showing considerable variability at the CSN1S1 gene in different geographical areas, and this needs further study to establish the relationship between the utility of goat germplasm in relation to human food, habit and migration. Analysis of geographical and molecular diversity along with food and habit (including other cultural diversity) will establish a new horizon for conserving the goat breeds.

The effects of $\alpha_S1$-casein polymorphism on milk yield and composition, micelle structure, renneting properties and cheese yield have been thoroughly studied in French breeds (Remeuf, 1993; Grosclaude et al., 1994; Mahé et al., 1994; Vassal et al., 1994; Barbieri et al., 1995; Ricordeau et al., 1996; Martin et al., 1999; Ricordeau et al., 2000). There is evidence that goats associated with the high content (A or B allele) of $\alpha_S1$-casein produce milk characterised by a significantly high percentage of protein, fat, total calcium, better curd-firming time, curd firmness and cheese yield compared with goats homozygous for alleles associated with a low or intermediate content (E or F allele). However, in our earlier reports, we had observed that $\alpha_S1$-casein genotype had a significant effect on protein content and calcium content in Jamunapari, Barbari, Marwari, Jakhraana and Sirohi goat breeds (Prakash et al., 2002; Rout et al., 2004a). The benefit from using the information on $\alpha_S1$-casein genotype in a selection programme for dairy goats will improve the protein content (Sanchez et al., 2005). The present study and previous reports on Indian goats also provide a clear indication that protein content can be improved by selecting the $\alpha_S1$-casein genotype. Therefore, genotyping at CSN1S1 locus should be carried out for better cheese yield and a genetic improvement programme.

Acknowledgements

The authors are grateful to the Director, CIRG for providing necessary facilities to carry out the work.

References


Boulanger A, Grosclaude F and Mahe MF 1984. $\alpha_S1$ and $\alpha_S2$ casein polymorphism in goats. Genetics Selection Evolution 16, 157–175.


