

## Genetic polymorphism of the kappa-casein gene in Brazilian cattle

A.L.S. Azevedo<sup>1,2</sup>, C.S. Nascimento<sup>2</sup>, R.S. Steinberg, M.R.S. Carvalho<sup>3</sup>,  
M.G.C.D. Peixoto<sup>1</sup>, R.L. Teodoro<sup>1</sup>, R.S. Verneque<sup>1</sup>, S.E.F. Guimarães<sup>2</sup> and  
M.A. Machado<sup>1</sup>

<sup>1</sup>Embrapa Gado de Leite, Juiz de Fora, MG, Brasil

<sup>2</sup>Departamento de Zootecnia, Universidade Federal de Viçosa, Viçosa, MG, Brasil

<sup>3</sup>Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais,  
Belo Horizonte, MG, Brasil

Corresponding author: M.A. Machado  
E-mail: machado@cnppl.embrapa.br

Genet. Mol. Res. 7 (3): 623-630 (2008)  
Received January 8, 2008  
Accepted June 20, 2008  
Published July 15, 2008

**ABSTRACT.** Frequencies of  $\kappa$ -casein gene alleles were determined in 1316 animals from the Brazilian *Bos indicus* genetic groups (Sindhi cows, Gyr sires, Gyr cows, Guzerat sires, Guzerat cows, Nellore sires, and Gyr x Holstein crossbreds) by means of polymerase chain reaction-restriction fragment length polymorphism analysis using two independent restriction nucleases (*Hinf*I and *Hae*III). The genotyping of  $\kappa$ -casein alleles (A and B) is of practical importance, since the B allele is found to correlate with commercially valuable parameters of cheese yielding efficiency. The frequencies of the B allele of  $\kappa$ -casein among breeds ranged from 0.01 to 0.30. The Sindhi breed had the highest frequency for the B allele (0.30), while the frequencies of this allele in other breeds ranged from 0.01 to 0.18. A wide variation in the B allele frequency among *B. indicus* breeds was

found suggesting that molecular selection for animals carrying the B allele could impact breeding programs for dairy production.

**Key words:** Zebu cattle; Casein; Candidate genes;  $\kappa$ -casein; Molecular markers

## INTRODUCTION

Caseins are milk proteins secreted by mammary gland cells. They constitute about 78-82% of bovine milk proteins and are subdivided into four main groups:  $\alpha_{s1}$ -casein,  $\alpha_{s2}$ -casein,  $\beta$ -casein, and  $\kappa$ -casein (Eigel et al., 1984). These proteins and their genetic variants have been extensively studied, and reported as an important factor associated with lactation performance, milk composition and cheese yield efficiency (Grosclaude, 1988; Aleandri et al., 1990). The casein genes are tightly linked and inherited as a cluster so they have a potential value and can play an important role in marker-assisted selection for milk traits (Lien and Rogne, 1993). They are located on chromosome 6 within a 200-kb fragment in the order  $\alpha_{s1}$ ,  $\beta$ ,  $\alpha_{s2}$  and  $\kappa$ . The  $\alpha_{s1}$ ,  $\beta$ , and  $\alpha_{s2}$ -casein genes are the most closely linked and form an evolutionarily related family, whereas the  $\kappa$ -casein gene is at least 70 kb away from them (Ferretti et al., 1990).

Kappa-casein ( $\kappa$ -Cn) constitutes approximately 12% of the casein and is a constituent of bovine milk. The gene is located on chromosome 6q31. The overall length of the  $\kappa$ -Cn gene is close to 13 kb, but most of the coding sequences for the mature  $\kappa$ -Cn protein are contained in the fourth exon.  $\kappa$ -Cn is considerably different from other caseins in structure and other properties. It is highly homologous to the  $\gamma$  chain of fibrinogen and fulfills a similar function, while serving as a stabilizing factor during the formation of micelle structure in curdling milk (Fiat and Jolles, 1989).

The  $\kappa$ -Cn gene has been broadly studied due to its influence on the manufacturing properties of milk. To date, nine variants have been described: A, B, C, E, F, G, H, I, and A<sub>1</sub>, the most frequent being the A and B alleles (Prinzenberg et al., 1999). The A and B variants differ in the amino acids 136 and 148. At position 136, threonine is replaced by isoleucine, while at position 148, aspartic acid is replaced by alanine, for A and B, respectively (Alexander et al., 1988). These two point mutations are located relatively close to several glycosylation sites and probably affect the structure of the protein and glycosylation patterns (Fox, 1992). The B allele was found to be associated with thermal resistance, shorter coagulation time, better curdles and micelles of different sizes, which are preferable in cheese making (Schaar et al., 1985; Marzali and Ng-Kwai-Hang, 1986). The cheese yield from cows with genotype BB is 10% higher compared to AA cows (Marzali and Ng-Kwai-Hang, 1986). The B allele of  $\kappa$ -Cn not only promotes an increase in cheese yield and improves cheese quality, but also correlates with other valuable parameters of milk productivity (protein content and milk yield). The effect of the  $\kappa$ -Cn B allele on the protein content of milk has also been reported in many studies involving Holstein Friesian breed cattle (Mao et al., 1992; Freyer et al., 1999; Gladyr et al., 2000; Tsiaras et al., 2005). This may be explained by the position of amino acid substitutions caused by mutations in exon IV of the  $\kappa$ -Cn gene.

Selection for the B allele of  $\kappa$ -Cn gene is integrated into cattle breeding programs in many countries. Nevertheless, most of the studies have been done with Taurine breeds and little is known about this gene in *indicus* breeds. Knowledge of the  $\kappa$ -Cn alleles frequency distribution in *indicus* breeds is of major importance for cattle breeding in tropical countries. We estimated the allele frequencies of  $\kappa$ -Cn alleles in Gyr, Guzarat, Nellore and Sindhi breeds.

## MATERIAL AND METHODS

### DNA samples

For DNA extraction, blood and semen were collected from pure and crossbred Zebu cattle (55 Sindhi cows, 73 Gyr x Holstein F1 cows, 182 Gyr sires, 564 Gyr cows, 69 Guzerat sires, 312 Guzerat cows, and 61 Nellore sires). DNA was extracted from the leukocytes or sperm cells with saline precipitation after protein removal. The quality and concentration of the DNA samples were determined with a GeneQuant Pro<sup>®</sup> spectrophotometer (Amersham Biosciences).

### Genotyping of $\kappa$ -Cn alleles

Genotypes of the A, B, and E alleles of the  $\kappa$ -Cn gene were determined using the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) technique. A 453-bp fragment of the exon IV was amplified with PCR using the primers CASK-F: 5'-TGTGCTGAGTAGGTATCCTAGTTATGG-3' and CASK-R: 5'-GCGTTGTCTTCTTTGATGTCTCCTTAG-3'. For this purpose, 60 ng genomic DNA was added to 25  $\mu$ L PCR mix containing PCR buffer, 2.5 mM MgCl<sub>2</sub>, 100  $\mu$ M dNTPs, 0.5  $\mu$ M of each of the primers and 1 U *Taq* DNA polymerase. Samples were amplified for 35 cycles: denaturation at 94°C for 60 s, annealing at 60°C for 45 s and extension at 72°C for 80 s, with a final extension step at 72°C for 10 min and storage at 4°C until be used.

PCR products were loaded onto 5% native polyacrylamide gel to evaluate amplification of the DNA samples prior to the digestion step. A total of 10  $\mu$ L of the PCR product was digested for 3 h at 37°C with 1 U *Hinf*I and 1 U *Hae*III restriction enzymes, in two separate reactions. The restriction fragments were separated on 8% native polyacrylamide gels.

### Statistical analysis

Estimates of the allele frequencies of  $\kappa$ -Cn were determined by direct count of the alleles divided by  $2n$  and the standard error of the allele frequencies calculated as:

$$\sqrt{\frac{q(1-q)}{2n}}, \quad (\text{Equation 1})$$

where  $q$  is the allele frequency and  $n$  the total number of samples tested (Weir, 1996).

The probability of Hardy-Weinberg equilibrium associated with the observed genotypic frequencies was obtained using the  $\chi^2$  test for each breed composition and the exact probability test (Haldane, 1954).

The observed heterozygosity was calculated by directly determining the frequency of heterozygosis:

$$Ho = \sum_{i=1}^n \frac{Hi}{n} \quad (\text{Equation 2})$$

where  $Hi$  is the number of heterozygous individuals and  $n$  is the total number of individuals analyzed.

The expected heterozygosity was calculated based on the observed allelic frequencies:

$$He = 1 - \sum_{i=1}^n p_i^2 \quad (\text{Equation 3})$$

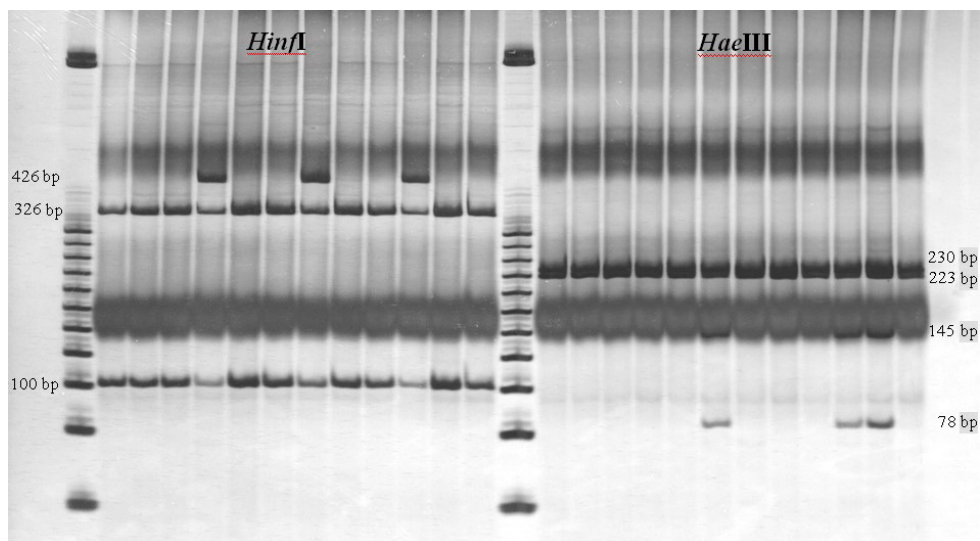
where  $p_i$  is the frequency of allele  $i$  and  $n$  is the number of alleles.

The statistical program TFPGA - Tools for Population Genetic Analysis (Miller, 1997) was used in the estimation of these parameters.

## RESULTS AND DISCUSSION

### Amplification

PCR amplification of the 453-bp fragment of the  $\kappa$ -Cn gene produced a clear and distinct band on the gel, which is of major importance in the digestion step since the presence of PCR artifacts could compromise the scoring of the fragments. Digestion of the PCR products with *Hinf*I generated four fragments: 426, 326, 100, and 27 bp, and digestion with *Hae*III generated four fragments: 230, 223, 145, and 78 bp (Figure 1). The 27-bp fragment generated by *Hinf*I could not be directly detected since it migrated very fast on the gel. Although nine alleles could be detected for the  $\kappa$ -Cn gene, we analyzed only the most common alleles (A, B and E). One important advantage of the detection of  $\kappa$ -Cn variants at the DNA level is the fact that sires can also be genotyped, whereas protein markers obtained from milk can be typed only for cows.



**Figure 1.** Silver-stained native polyacrylamide gel showing  $\kappa$ -Cn gene polymorphisms revealed by polymerase chain reaction-restriction fragment length polymorphism in 12 bovines.

### Allele and genotype frequencies

Allele and genotype frequencies for 1316 animals from the Gyr, Gyr x Holstein F1, Guzerat, Nellore, and Sindhi breeds are shown in Table 1. For all breeds, the A allele was the most frequent, with values ranging from 0.7 to 0.99. These findings are similar to those of Baker and Manwell (1980), who reported a much higher frequency of the A variant detected with protein markers.

**Table 1.** Allele and genotype frequency of  $\kappa$ -Cn alleles A, B and E in Gyr, Guzerat, Nellore, and Sindhi breeds according to the genetic group.

Breed or genetic group	Sample	N	Allele frequency				Genotype frequency			
			A	B	E	SEM	AA	AB	BB	AE
Sindhi	Cows	55	0.70	0.30	0.00	± 0.04	0.45	0.50	0.05	0.00
Gyr x Holstein F1	Cows	73	0.82	0.13	0.05	± 0.03	0.64	0.26	0.00	0.10
Gyr	Sires	182	0.94	0.06	0.00	± 0.04	0.88	0.12	0.00	0.00
	Cows (Herd 1)	414	0.90	0.10	0.00	± 0.01	0.81	0.18	0.01	0.00
	Cows (Herd 2)	150	0.99	0.01	0.00	± 0.01	0.97	0.03	0.00	0.00
	Combined	746	0.93	0.07	0.00	± 0.01	0.86	0.13	0.01	0.00
Guzerat	Sires	69	0.86	0.14	0.00	± 0.03	0.74	0.25	0.01	0.00
	Cows	312	0.82	0.18	0.00	± 0.01	0.68	0.27	0.05	0.00
	Combined	381	0.83	0.17	0.00	± 0.01	0.69	0.26	0.05	0.00
Nellore	Sires	61	0.82	0.18	0.00	± 0.03	0.67	0.30	0.03	0.00

SEM = standard error of the mean.

The E allele was only found in the Gyr x Holstein F1 cows, with a very low frequency of 0.05. This allele is found almost exclusively in breeds at low frequencies. Prinzenberg et al. (1999) investigated various Taurine breeds and two *indicus* crossbreeds and detected the presence of the E allele in Ayshire, Holstein and Red Cattle breeds (allele frequency ranged from 0.08 to 0.17). The E allele in the Gyr x Holstein F1 most probably originated from the Holstein breed.

The B allele was found in all breeds, although, present at low frequencies, ranging from 0.01 to 0.3. Similar results were reported in the few publications on *indicus* breeds. Kemenes et al. (1999) investigated 83 animals and reported a frequency of 0.07 in the Gyr breed for the B allele. In the present study, a total of 746 Gyr cattle were genotyped, including two herds of cows and sires participating in a progeny test program for milk production traits; we also found a combined frequency of 0.07 for the B allele. We found a large difference in the frequency of the B allele between Gyr herd 1 (0.10) and Gyr herd 2 (0.01) (Table 1). These herds have been independently selected for milk production over many years. Founder effects of the sires used in artificial insemination in each herd could explain the difference in the frequency of the B allele between these two populations.

The frequency of the B allele found in the Gyr x Holstein F1 cows was 0.13. This low frequency was expected since the frequency of this allele in both founder breeds was also low (Ng-Kwai-Hang et al., 1990; Famula and Medrano, 1994). Tsiaras et al. (2005) found a frequency of 0.06 for this allele in a sample of 278 Holsteins.

A total of 381 Guzerat animals were analyzed and a combined frequency for the B allele of 0.17 was found. The frequency of the B allele was 0.14 for the sires and 0.18 for the cows. Similar results were reported by Silva and Del Lama (1997) who analyzed protein profiles of milk samples from 205 animals of the Guzerat breed, and found a frequency of 0.19 for the variant B.

Biase et al. (2005) genotyped 408 Nellore animals and found a frequency of 0.09 for the B allele, while Silva and Del Lama (1997) found a frequency of 0.03 in this breed. In the present study, a much higher frequency was reported for the B allele in the Nellore breed (0.18) compared to previous reports. There is a great variation in the  $\kappa$ -Cn gene in this breed which could be by differences in selection strategies within the enormous Nellore population in Brazil.

The highest frequency for the B allele was found in the Sindhi breed (0.30), which had a much higher frequency than the other breeds (Table 1). Silva and Del Lama (1997) analyzed milk proteins from 22 Sindhi cows and found a frequency of 0.34 for the variant B, very close to the frequency that we found.

### Hardy-Weinberg equilibrium

Tests of departure from Hardy-Weinberg proportions are frequently made to check on random mating in populations, and the deviations from expectation are used to estimate inbreeding coefficients. A population is considered to be in Hardy-Weinberg equilibrium when the allelic ( $p$  and  $q$ ) and genotypic ( $p^2$ ,  $2pq$  and  $q^2$ ) frequencies remain constant from one generation to the next, due to a random association of gametes. According to Tambasco et al. (2000) and Vasconcellos et al. (2003), some events, such as an accumulation of some genotypes, subdivision of the population, mutation, selection, migration, or endogamy, can result in a state of disequilibrium within the population. All breeds and genetic groups were found to be in Hardy-Weinberg equilibrium (Table 2), suggesting that the  $\kappa$ -Cn gene was not influenced by selection.

**Table 2.** Test for Hardy-Weinberg (HW) equilibrium of  $\kappa$ -Cn in Gyr, Guzerat, Nellore, and Sindhi breeds according to the genetic group.

Breed or genetic group	Sample	P (HW)
Sindhi	Cows	0.34
Gyr	Sires	1.00
	Cows	1.00
	Combined	1.00
Guzerat	Sires	1.00
	Cows	0.06
	Combined	0.08
Nellore	Sires	1.00

P = exact probability (Haldane, 1954).

### CONCLUSIONS

In Brazil, milk producers have only recently started to pay attention to the manufacturing properties of milk, because some industries are beginning to differentially remunerate the milk according to its composition in terms of fat and protein percentage as well as cheese-making properties. The wide variation found for  $\kappa$ -Cn gene in the Brazilian *indicus* breeds that we found probably reflects the overall existing variation in milk constituents in these breeds. In addition, the increasing demand by the industry for milk constituents should greatly impact on dairy cattle breeding strategies in Brazil.

Because of the effects of  $\kappa$ -Cn genetic variants on cheese yield, selection of animals

with the favorable  $\kappa$ -Cn B allele is considerable. As there is a low frequency of this allele in the Brazilian *indicus* breeds, selection directed by molecular markers could positively impact on dairy production.

## ACKNOWLEDGMENTS

We thank ABCGIL (Brazilian Gyr Dairy Cattle Association) for the Gyr blood and semen samples and CBMG (Brazilian Center for Guzerat Improvement) for the Guzerat blood and semen. C.S. Nascimento was the recipient a scholarship from FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais). Research supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and Embrapa PRODETAB (Projeto de Apoio ao Desenvolvimento de Tecnologias Agropecuárias para o Brasil).

## REFERENCES

- Aleandri RL, Buttazoni G, Schneider JC, Caroli A, et al. (1990). The effects of milk protein polymorphisms on milk components and cheese-producing ability. *J. Dairy Sci.* 73: 241-255.
- Alexander LJ, Stewart AF, Mackinlay AG, Kapelinskaya TV, et al. (1988). Isolation and characterization of the bovine kappa-casein gene. *Eur. J. Biochem.* 178: 395-401.
- Baker CMA and Manwell C (1980). Chemical classification of cattle. *Anim. Blood Groups Biochem. Genet.* 11: 127-150.
- Biase FH, Garnero A Del V, Bezerra LF and Rosa JMA (2005). Analysis of restriction fragment length polymorphism in the kappa-casein gene related to weight expected progeny difference in Nellore cattle. *Genet. Mol. Biol.* 28: 84-87.
- Eigel WN, Butler JE, Ernstrom CA and Farrell HM (1984). Nomenclature of proteins of cow's milk: Fifth revision. *J. Dairy Sci.* 67: 1599-1631.
- Famula TR and Medrano JF (1994). Estimation of genotype effects for milk proteins with animal and sire transmitting ability models. *J. Dairy Sci.* 77: 3153-3162.
- Ferretti L, Leone P and Sgaramella V (1990). Long range restriction analysis of the bovine casein genes. *Nucleic Acids Res.* 18: 6829-6833.
- Fiat AM and Jolles P (1989). Caseins of various origins and biologically active casein peptides and oligosaccharides: structural and physiological aspects. *Mol. Cell Biochem.* 87: 5-30.
- Fox PF (1992). *Advanced Dairy Chemistry Proteins*. Vol. 1. Elsevier Science Publisher, London.
- Freyer G, Liu Z, Erhardt G and Panicke L (1999). Casein polymorphism and between milk production traits. *J. Anim. Breed Genet.* 116: 87-97.
- Gladyr EA, Zinoveva NA, Marzanov NS and Brem G (2000). The use of beta-lactoglobulin and kappa-casein genes as the genetic markers for cattle. In: *Biotechnology in Crop Production, Livestock Farming, and Veterinary*. Proceedings of the II International Scientific Conference, Moscow, 86-88.
- Grosclaude F (1988). Le polymorphisme génétique des principales lactoprotéines bovines. Relations avec la quantité, la composition et les aptitudes fromagères du lait. *INRA Prod. Anim.* 1: 5-7.
- Haldane JBS (1954). An exact test for randomness of mating. *J. Genet.* 52: 631-635.
- Kemenes PA, Regitano LCA, Rosa AJM, Packer IU, et al. (1999).  $\kappa$ -Casein,  $\beta$ -lactoglobulin and growth hormone allele frequencies and genetic distances in Nelore, Gyr, Guzerá, Caracu, Charolais, Canchim and Santa Gertrudis cattle. *Genet. Mol. Biol.* 22: 539-541.
- Lien S and Rogne S (1993). Bovine casein haplotypes: number, frequencies and applicability as genetic markers. *Anim. Genet.* 24: 373-376.
- Mao IL, Buttazoni LG and Aleandri R (1992). Effects of polymorphic milk protein genes on milk yield and composition traits in Holstein cattle. *Acta Agric. Scand.* 42: 1-7.
- Marziali AS and Ng-Kwai-Hang KF (1986). Effects of milk composition and genetic polymorphism on cheese composition. *J. Dairy Sci.* 69: 2533-2542.
- Miller MP (1997). Tools for population genetics analyses (TFPGA): a Windows program for the analysis of allozyme and molecular population genetic data. Available at <http://www.marksgeneticssoftware.net/tfpga.htm>.
- Ng-Kwai-Hang KF, Monardes HG and Hayes JF (1990). Association between genetic polymorphism of milk proteins and production traits during three lactations. *J. Dairy Sci.* 73: 3414-3420.

- Prinzenberg EM, Krause I and 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). *Anim. Biotechnol.* 10: 49-62.
- Schaar J, Hansson B and Pettersson H (1985). Effects of genetic variants of  $\kappa$ -casein and beta lactoglobulin on cheese-making. *J. Dairy Res.* 52: 429-437.
- Silva IT and Del Lama M (1997). Milk protein polymorphisms in Brazilian Zebu cattle. *Braz. J. Genet.* 20: 625-630.
- Tambasco DD, Alencar MM, Coutinho LL and Tambasco AJ (2000). Caracterização molecular de animais da raça Nelore utilizando microssatélites e genes candidatos. *Rev. Bras. Zoot.* 29: 1044-1049.
- Tsiaras AM, Bargouli GG, Banos G and Boscós CM (2005). Effect of kappa-casein and beta-lactoglobulin loci on milk production traits and reproductive performance of Holstein cows. *J. Dairy Sci.* 88: 327-334.
- Vasconcellos LPMK, Tambasco DD, Pereira AP and Pereira AP (2003). Genetic characterization of Aberdeen Angus cattle using molecular markers. *Genet. Mol. Biol.* 26: 133-137.
- Weir BS (1996). *Genetic Data Analysis: Methods for Discrete Population Genetic Data*. 2nd edn. Sinauer Associates, Sunderland.