

## Genetic polymorphism of milk protein loci in Argentinian Holstein cattle

Adriana Gloria Bonvillani, Miguel Angel Di Renzo and Iván Nicolás Tiranti

### Abstract

Some alleles of milk protein loci are associated with superior cheese production characteristics. The genetic polymorphism of the milk protein loci  $\alpha_{s1}$ -casein,  $\beta$ -casein,  $\kappa$ -casein and  $\beta$ -lactoglobulin was examined in Argentinian Holstein cattle. Samples from 12 herds of four regions of Córdoba were analyzed by starch gel electrophoresis. The  $\chi^2$  test was used to assess whether the populations were in Hardy-Weinberg equilibrium. Genotypic diversity was analyzed by the Shannon-Weaver index. The observed genotypic frequencies were analyzed by Hedrick's genetic identity and the genetic distance of Balakrishnan and Sanghvi. The allelic and genotypic frequencies were similar to those of other Holstein populations. The genotypic frequencies of the  $\alpha_{s1}$ -casein and  $\beta$ -casein loci were in equilibrium, whereas in some populations the  $\kappa$ -casein and  $\beta$ -lactoglobulin loci were not. According to the Shannon-Weaver index the total genetic diversity within each herd was greater than 96%. The high values of identity agreed with the low genetic distances among populations. We conclude that there is extensive genetic homogeneity in Holstein cattle in Córdoba Province and that it would be feasible to select for B alleles at the  $\kappa$ -casein and  $\beta$ -lactoglobulin loci in order to improve the quality of milk available for cheese manufacturing.

### INTRODUCTION

The milk proteins  $\alpha_{s1}$ -casein,  $\beta$ -casein,  $\kappa$ -casein and  $\beta$ -lactoglobulin are determined by the loci CASA1, CASB, CASK and LGB, respectively. Each of these loci has different alleles coding for several genetic variants of the corresponding protein. The  $\alpha_{s1}$ -casein locus has three main alleles (A, B and C). Allele A is found in Red Danish cattle and also is typical of the Holstein breed, but at very low frequencies. Allele B predominates in European breeds, whereas allele C is frequent in zebu breeds, but absent in Holsteins. An additional variant D is reported to occur in French and Italian breeds (Thompson *et al.*, 1962; Li and Gaunt, 1972). The most widely distributed alleles of the  $\beta$ -casein locus are the A and B alleles, the former being present at higher frequencies than the latter. The C variant of the  $\alpha_{s1}$ -casein locus is found in some European breeds but not in Holsteins (Aschaffenburg and Drewry, 1955; Li and Gaunt, 1972). The  $\kappa$ -casein locus has only two alleles, A and B, with  $\kappa$ -casein A being more frequent than  $\kappa$ -casein B in almost all breeds (Neelin, 1964; Li and Gaunt, 1972). Alleles A and B of the  $\beta$ -lactoglobulin locus are universally present, with a predominance of B in most breeds. Two other alleles, C and D, are typical of Jersey and French cattle, respectively (Aschaffenburg and Thymann, 1965; Kiddy *et al.*, 1972; Li and Gaunt, 1972). Numerous studies have investigated the genetic polymorphisms of milk protein loci in a variety of cattle breeds (McLean *et al.*, 1984; Ng-Kwai-Hang *et al.*, 1984, 1986; Tejedor *et al.*, 1986; Grosclaude, 1988; Van Eenennaam and Medrano, 1991; Silva and Del Lama, 1997). In Argentina, such polymorphisms were first studied by Poli and Antonini (1991) in

Creole and Argentinian Holstein cattle, the latter being the most important dairy breed in this country.

Several studies that have examined the loci involved in the genetic control of milk proteins in cattle have concluded that the alleles of these loci may be used as genetic markers of quantitative trait loci (QTL) for milk yield, composition and quality. For the  $\alpha_{s1}$ -casein locus, the presence of the B allele increases milk and fat yields as well as  $\beta$ -casein and  $\kappa$ -casein contents. Similarly, the B allele is associated with greater fat and protein yields in the  $\beta$ -casein locus. The B allele of the  $\kappa$ -casein locus also leads to superior milk and protein yields. For the  $\beta$ -lactoglobulin locus, allele A is associated with higher milk and protein yields, whereas the B allele is associated with fat and protein concentrations (McLean *et al.*, 1984; Ng-Kwai-Hang *et al.*, 1986; Van Eenennaam and Medrano, 1991; Bovenhuis *et al.*, 1992). Milk from cattle with B alleles for the loci  $\beta$ -casein,  $\kappa$ -casein and  $\beta$ -lactoglobulin is better for cheese manufacturing because of its greater content of dry matter which reflects the high fat and protein content (Marziali and Ng-Kwai-Hang, 1986; Morini *et al.*, 1982).

Our objective was to analyze the genetic polymorphism and the genetic variability distribution of milk protein loci in Holstein cattle of Córdoba Province, Argentina.

### MATERIAL AND METHODS

Thirty samples per herd were collected from 12 herds from the four main dairy cattle regions in Córdoba Province (Table I). Identification of the genetic variants was performed by starch gel electrophoresis at alkaline pH (Kim and Bird, 1972). Allelic frequencies were determined

by direct counting. The  $\chi^2$  test was used to check whether the populations were in Hardy-Weinberg equilibrium.

The genotypic frequencies were used in subsequent analyses because of the close linkage among the casein loci (Threadgill and Womack, 1990). The Shannon-Weaver diversity index ( $I'$ ) was used to quantify the genetic diversity for each locus within and among herds and regions. Lewontin (1972) adapted this ecological measure to calculate genetic diversity in human populations. The milk protein genotypic frequencies were analyzed using the diversity index  $I'$ , defined as:

$$I' = - \sum_{i=1}^n p_i \log p_i \text{ with } \sum_{i=1}^n p_i = 1$$

where  $p_i$  is the relative genotypic frequency at the  $i^{\text{th}}$  locus from each herd and region and  $n$  is the number of genotypes. The partitioning of the total intra- and intergroup diversity was based on the mean diversities of the herds ( $I'_H$ ) or regions ( $I'_S$ ); the diversity  $I'_T$  was calculated from the sum of  $p_i$ .

The observed genotypic frequencies were analyzed by the genetic identity measures of Hedrick (1971) and by the distance coefficient of Balakrishnan and Sanghvi (1968). These indices were used to calculate the identity and distance among herds and regions. The results were summarized in a matrix derived using the unweighted pair group method, with arithmetic means clustering or UPGMA (Sneath and Sokal, 1973).

## RESULTS AND DISCUSSION

The  $\alpha_{s1}$ -casein locus had no AA or CC genotypes. The allelic form  $\alpha_{s1}$ -casein A was not found in this study, al-

though it is commonly present at low frequencies in this breed. The B variant predominated at this locus (Table I). The most common genotype at the  $\beta$ -casein locus was AA (Table I). AB heterozygosity occurred at a low frequency; there were no BB animals. The  $\kappa$ -casein A variant was the most important variant at this locus (Table I) and the  $\kappa$ -casein AA genotype was more frequent than the AB or BB genotype. Approximately half of the cows were heterozygous (AB) at the  $\beta$ -lactoglobulin locus, with a slight predominance of the  $\beta$ -lactoglobulin B allele (Table I). The allelic and genotypic frequencies in Argentinian Holstein cattle from Córdoba were similar to those for other studies of this breed in other countries (Aschaffenburg and Drewry, 1955; Li and Gaunt, 1972; Ng-Kwai-Hang *et al.*, 1984; Poli and Antonini, 1991; Van Eenennaam and Medrano, 1991). High frequencies for the  $\alpha_{s1}$ -casein B and  $\beta$ -casein A alleles (close to the fixation limit) have been reported previously for Holstein cattle (Ng-Kwai-Hang *et al.*, 1986; Van Eenennaam and Medrano, 1991). Since these alleles are associated with milk and fat yields, a long-term, intensive selection aimed at improving milk production might have indirectly increased their frequencies (Ng-Kwai-Hang *et al.*, 1986). The predominance of the  $\kappa$ -casein A allele cannot be explained by selection since it is the  $\kappa$ -casein B allele that is associated with high fat and protein contents. A possible cause may be the linkage present among casein alleles since the most frequent haplotype is  $\alpha_{s1}$ -casein B -  $\beta$ -casein A -  $\kappa$ -casein A.

The genotypic frequencies observed in most of the herds and regions were within the Hardy-Weinberg expectations for the  $\alpha_{s1}$ -casein and  $\beta$ -casein loci (Table I) in agreement with other studies (Li and Gaunt, 1972; Ng-Kwai-Hang *et al.*, 1984; Poli and Antonini 1991). There was a

**Table I** - Allelic frequencies at the  $\alpha_{s1}$ -casein,  $\beta$ -casein,  $\kappa$ -casein and  $\beta$ -lactoglobulin loci in Holstein herd milk together with the  $\chi^2$  test for Hardy-Weinberg equilibrium.

Herds	$\alpha_{s1}$ -casein		$\beta$ -casein		$\kappa$ -casein		$\beta$ -lactoglobulin	
	B	$\chi^2$	A	$\chi^2$	A	$\chi^2$	A	$\chi^2$
Coronel Moldes 1	1.00	-	0.98	N.S.	0.73	N.S.	0.52	N.S.
Coronel Moldes 2	0.95	N.S.	0.97	N.S.	0.75	**	0.35	N.S.
Coronel Moldes 3	0.98	N.S.	0.97	N.S.	0.80	*	0.47	N.S.
Coronel Moldes 4	0.97	N.S.	0.95	N.S.	0.75	*	0.42	N.S.
Coronel Moldes 5	0.95	N.S.	0.95	N.S.	0.78	N.S.	0.40	*
Villa María 6	0.98	N.S.	0.97	N.S.	0.77	*	0.38	N.S.
Villa María 7	0.97	N.S.	0.98	N.S.	0.78	**	0.41	N.S.
Devoto 8	0.97	N.S.	0.97	N.S.	0.77	N.S.	0.45	N.S.
Devoto 9	0.98	N.S.	0.98	N.S.	0.78	N.S.	0.48	*
La Carlota 10	0.97	N.S.	1.00	-	0.73	N.S.	0.39	N.S.
La Carlota 11	0.95	N.S.	0.93	N.S.	0.75	N.S.	0.45	N.S.
La Carlota 12	0.95	N.S.	0.97	N.S.	0.77	*	0.48	N.S.
<b>Regions</b>								
Coronel Moldes	0.97	N.S.	0.96	N.S.	0.76	**	0.43	N.S.
Villa María	0.98	N.S.	0.98	N.S.	0.78	**	0.39	N.S.
Devoto	0.98	N.S.	0.98	N.S.	0.78	N.S.	0.47	**
La Carlota	0.96	N.S.	0.97	N.S.	0.75	**	0.44	N.S.
<b>Total</b>	0.97	N.S.	0.97	N.S.	0.76	**	0.43	N.S.

N.S., Nonsignificant. \* $P < 0.05$ . \*\* $P < 0.01$ .

departure from equilibrium for the  $\kappa$ -casein and  $\beta$ -lactoglobulin loci in some populations (Table I). This significant deviation from equilibrium at the  $\kappa$ -casein locus in six herds from Coronel Moldes, Villa María and La Carlota resulted from a greater proportion of homozygous AA cows in these populations which may reflect the linkage noted above. The herds Coronel Moldes 5 and Devoto 9 and the Devoto region did not follow the expected genotypic proportions at the  $\beta$ -lactoglobulin locus because of a higher frequency of homozygous BB in the former and an excess of heterozygous animals in the latter two populations. This predominance of AB individuals was reported previously by Van Eenennaam and Medrano (1991) and was explained as a difference in the selective value of one of the alleles at the locus. This observation could also have resulted from the selection pressure that has been applied in dairy cattle for increased milk and fat production. These traits are associated with A and B respectively alleles at the  $\beta$ -lactoglobulin locus. Therefore, the selection could have indirectly maintained both alleles in the population and increased the heterozygous frequency.

The Shannon-Weaver index calculated for each locus was highest for the  $\beta$ -lactoglobulin locus with values of 1.39 in the herds and 1.42 among the regions; the total  $I'$  value was 1.46. The  $I'$  value for  $\kappa$ -casein was 1.25 in the herds and at the regions, with a total  $I'$  of 1.26. For the  $\alpha_{s1}$ -casein and  $\beta$ -casein loci, the value was 0.33 in the herds and at the regions, with a total  $I'$  of 0.34 for both genetic markers. This index showed that most of the genetic variation was at the  $\beta$ -lactoglobulin and  $\kappa$ -casein loci, each of which had three genotypes (AA, AB and BB). The former locus had greater genetic variation because of a higher proportion of heterozygotes that maintain the two alleles in equally frequent proportions in the populations. The latter locus showed less variation because of the predominance of AA genotype and the A allele. The genetic diversity for  $\alpha_{s1}$ -casein and  $\beta$ -casein loci was practically null because the genotypes  $\alpha_{s1}$ -casein BB and  $\beta$ -casein AA had frequencies greater than 0.90.

According to the apportionment of the Shannon-Weaver index for each locus over 90% of the genetic variation occurred within each herd (Table II). The contribution to the total diversity of subdivisions among herds within and among regions was very low. Considering this index as the mean of the four loci analyzed (Table II), over 96% of the genetic diversity occurred within each herd. Variation

resulting from subdivisions among herds within and among regions contributed 0.4 and 2.8%, respectively, to the total diversity. Most of the genetic diversity was within each herd and there were no genetic differences among herds or regions.

The high values of identity were congruent with the low values for genetic distances among herds (data not shown) and regions (Table III). The matrix obtained using the Hedrick index showed high values of genetic identity among herds and regions. The identity values oscillated from 0.950 to 0.998 in the herds and the identity among regions yielded values next to 1 (Table III). The Balakrishnan and Sanghvi coefficient showed genetic distances below 0.369 in the herds and the distances among regions were rather low (Table III).

The herds Coronel Moldes 4 and Devoto 8 had the lowest identity value (0.95) which agreed with the large distance between them (0.37). A similar relationship was not observed for the other extreme, where the highest identity was between herds Villa María 6 and 7 (0.99) and the lowest distance was between Coronel Moldes 5 and La Carlota 11 (0.028); the next lowest distance was between Villa María 6 and 7 (0.031). Among regions, the identity values were greater than 0.98 and the distances were lower than 0.14. Devoto region differed from the other three regions which had similar identity and distance values. The Hedrick identity index and the Balakrishnan and Sanghvi distance coefficient revealed great genetic similarity among herds and regions for the genetic markers analyzed.

The analyses used to evaluate genetic variability, genetic similarity and distance, gave similar results. Most of the genetic diversity at these loci occurred within each herd and there was no significant genetic differentiation among herds and among regions. Thus, none of the 12 herds nor the four regions could be identified solely by the frequencies of the four milk protein loci studied. There was genetic diversity among individuals but there was no genetic differentiation among populations. This genetic homogeneity of the population could be explained by the fact that there was a great genetic flow among herds and regions due to an intensive commercial interchange of individuals among populations. It also appeared that the bulls for artificial insemination were genetically different and that they were used at random in populations. These two causes maintain diversity among individuals but blur group identity among subdivisions. This leads us to conclude that this Ar-

**Table II** - Hierarchical Shannon-Weaver index of herds and regions.

Milk protein	$\alpha_{s1}$ -casein	$\beta$ -casein	$\kappa$ -casein	$\beta$ -lactoglobulin	Mean	
$I'_H/I'_T$	0.95	0.98	0.99	0.95	0.960	Within herds
$(I'_S - I'_H)/I'_T$	0.02	0.02	0.00	0.02	0.004	Among herds
$(I'_T - I'_S)/I'_T$	0.03	0.03	0.01	0.03	0.028	Among regions

$I'_H$ ,  $I'_S$  and  $I'_T$ : Herd, region and total diversity, respectively.

**Table III** - Matrix calculated by Hedrick identities (below the diagonal) and Balakrishnan and Sanghvi distances (above the diagonal) from observed genotypic frequencies of milk protein loci in Argentinian Holstein from different geographical regions of Córdoba.

	Coronel Moldes	Villa María	Devoto	La Carlota
Coronel Moldes	-	0.016	0.100	0.013
Villa María	0.998	-	0.138	0.045
Devoto	0.987	0.982	-	0.074
La Carlota	0.998	0.995	0.992	-

gentinian Holstein group of cattle from Córdoba constitutes a single population for the loci studied and that it has a panmictic behavior, which was confirmed by the Hardy-Weinberg equilibrium of the total genotypic frequencies (Table I).

According to the allelic frequencies determined in this study the alleles  $\alpha_{s1}$ -casein B and  $\beta$ -casein A were near the fixation limit and it would be difficult to change their frequencies. In contrast, the frequencies of  $\beta$ -lactoglobulin and  $\kappa$ -casein loci could be altered. The selection for B alleles at these loci could possibly improve the quality of milk for manufacturing processes. The frequency of  $\beta$ -lactoglobulin B was moderate and it would be possible to increase this variant within a short period of time. Although the  $\kappa$ -casein B allele frequency was low, it could also be changed. As these two loci are located on different chromosomes, linkage would not complicate selection favoring these alleles.

Of the total volume of milk (165 million liters/month) produced by Holstein cattle in Córdoba province in recent years, 60% was used in cheese manufacturing. The findings of this study suggest that it would be very useful to include the  $\beta$ -lactoglobulin and  $\kappa$ -casein loci as genetic markers for polygenic traits in traditional breeding programs designed to improve the quality of milk. Considering the homogeneity of this breed, it would be possible to establish the same selection program for all populations by using bulls with BB genotypes in conjunction with artificial insemination and the progressive replacement of AA cows with AB and BB individuals.

#### ACKNOWLEDGMENTS

This research was supported by the Secretaría de Ciencia y Técnica of the Universidad Nacional de Río Cuarto and the Consejo de Investigaciones Científicas y Técnicas Córdoba, Argentina. The authors are indebted to Prof. Iliana Martínez for reviewing the English of the manuscript, and thank Ing. Javier Turró for help in obtaining the milk samples.

#### RESUMO

Alguns alelos de locos de proteínas do leite associam-se com produção mais elevada de queijo. O polimorfismo genético

dos locos das proteínas do leite  $\alpha_{s1}$ -caseína,  $\beta$ -caseína,  $\kappa$ -caseína e  $\beta$ -lactoglobulina foi examinado em gado Holstein argentino. Amostras de 12 rebanhos de quatro regiões de Córdoba foram analisados por eletroforese em gel de amido. O teste do  $\chi^2$  foi usado para avaliar se as populações estavam em equilíbrio de Hardy-Weinberg. A diversidade genotípica foi analisada pelo índice de Shannon-Weaver. As frequências genotípicas observadas foram analisadas pela identidade genética de Hedrick e pela distância genética de Balakrishnan e Sanghvi. As frequências alélicas e genotípicas foram semelhantes às de outras populações de Holstein. As frequências genotípicas dos locos de  $\alpha_{s1}$ -caseína e  $\beta$ -caseína estavam em equilíbrio, enquanto que em algumas populações os locos de  $\kappa$ -caseína e  $\beta$ -lactoglobulina não estavam. De acordo com o índice de Shannon-Weaver a diversidade genética total dentro de cada rebanho foi maior que 96%. Os altos valores de identidade concordaram com as pequenas distâncias genéticas entre as populações. Concluímos que há extensa homogeneidade genética no gado Holstein na Província de Córdoba e que seria factível seleccionar os alelos B nos locos de  $\kappa$ -caseína e  $\beta$ -lactoglobulina de modo a melhorar a qualidade do leite disponível para a manufatura de queijo.

#### REFERENCES

- Aschaffenburg, R. and Drewry, J. (1955). Occurrence of different beta-lactoglobulins in cow's milk. *Nature* 176: 218-219.
- Aschaffenburg, R. and Thymann, M. (1965). Simultaneous phenotyping procedure for the principal proteins of cow's milk. *J. Dairy Sci.* 48: 1524-1526.
- Balakrishnan, V. and Sanghvi, L. (1968). Distance between populations on the basis of attribute data. *Biometrics* 24: 859-865.
- Bovenhuis, H., Van Arendonk, J. and Korver, S. (1992). Associations between milk protein polymorphisms and milk production traits. *J. Dairy Sci.* 75: 2549-2559.
- Grosclaude, F. (1988). Le polymorphisme génétique des principales lactoprotéines bovines. *INRA Prod. Anim.* 1: 5-17.
- Hedrick, P. (1971). A new approach to measuring genetic similarity. *Evolution* 25: 276-280.
- Kiddy, C., Rollins, R. and Zikakis, J. (1972). Discontinuous polyacrylamide electrophoresis for beta-lactoglobulins typing of cow's milk. *J. Dairy Sci.* 55: 1506-1507.
- Kim, G. and Bird, E. (1972). Effect of method of preparation on composition and electrophoretic properties of isoelectric casein. *J. Dairy Sci.* 55: 897-900.
- Lewontin, R. (1972). The apportionment of human diversity. *Evol. Biol.* 6: 381-398.
- Li, H. and Gaunt, S. (1972). A study of genetic polymorphisms of milk  $\beta$ -lactoglobulin,  $\alpha_{s1}$ -casein,  $\beta$ -casein, and  $\kappa$ -casein in five dairy breeds. *Biochem. Genet.* 6: 9-20.
- Marziali, A. and Ng-Kwai-Hang, K. (1986). Effects of milk composition and genetic polymorphism on cheese composition. *J. Dairy Sci.* 69: 2533-2542.
- McLean, D., Graham, B., Ponzoni, W. and McKenzie, H. (1984). Effects of milk protein genetic variants on milk yield and composition. *J. Dairy Res.* 51: 531-546.
- Morini, D., Castagneti, G., Chiavaric, C., Losi, G., Davoli, R. and Bosi, P. (1982). Prova di caseificazione con latte caratterizzato dalle varianti A e B della beta-attoglobulina nella produzione del formaggio parmigiano-reggiano. *Sci. Tecn. Latt. Cas.* 33: 475-492.
- Neelin, J. (1964). Variant of  $\kappa$ -casein revealed by improved starch gel electrophoresis. *J. Dairy Sci.* 47: 506-510.
- Ng-Kwai-Hang, K., Hayes, J., Moxley, J. and Monardes, G. (1984). Association of genetic variants of casein and milk serum proteins with milk, fat and protein production by dairy cattle. *J. Dairy Sci.* 67: 835-840.
- Ng-Kwai-Hang, K., Hayes, J., Moxley, J. and Monardes, G. (1986). Relationships between milk protein polymorphisms and major constituents in Holstein-Friesian cows. *J. Dairy Sci.* 69: 22-26.

- Poli, M. and Antonini, A.** (1991). Genetic structure of milk proteins in Argentinian Holstein and Argentinian Creole cattle. *Hereditas* 115: 177-182.
- Silva, I.V. and Del Lama, M.A.** (1997). Milk protein polymorphisms in Brazilian Zebu cattle. *Braz. J. Genet.* 20: 625-630.
- Sneath, P. and Sokal, R.** (1973). *Numerical Taxonomy*. W.H. Freeman and Co., San Francisco.
- Tejedor, T., Zarazaga, I. and Altarriba, J.** (1986). Consideraciones sobre la mejora genética de las características tecnológicas de la leche en ganado Holstein-Friesian. *ITEA* 66: 27-33.
- Thompson, M., Kiddy, C., Pepper, L. and Zittle, C.** (1962). Variations in the  $\alpha_s$ -casein of individual cow's milk. *Nature* 195: 1001-1002.
- Threadgill, D. and Womack, J.** (1990). Genomic analysis of the major bovine milk protein genes. *Nucleic Acids Res.* 18: 6935-6942.
- Van Eenennaam, A. and Medrano, J.** (1991). Milk protein polymorphisms in California dairy cattle. *J. Dairy Sci.* 74: 1730-1742.

(Received April 12, 1999)

