Journal of Agricultural Science, Cambridge (1998), **131**, 97–101. *Printed in the United Kingdom*
© 1998 *Cambridge University Press*

β-lactoglobulin variability in Argentinian Holstein cattle

A. G. BONVILLANI*, M. A. DI RENZO, A. MONTILLA AND I. N. TIRANTI

Facultad de Agronomía y Veterinaria, Universidad Nacional de Río Cuarto, 5800 Río Cuarto, Argentina

(*Revised MS received 19 December 1997*)

SUMMARY

To determine the genetic variability and population structure of the β-lactoglobulin (LGB) locus in the Argentinian Holstein breed, milk samples from 12 herds at four locations in Córdoba, Argentina, were analysed. Allelic and genotypic frequencies of LGB variants were compared by contingency tables. A G-test was applied to detect Hardy–Weinberg equilibrium. LGB A allele frequency was 0·43 and B allele frequency was 0·57. Allelic and genotypic frequencies indicated great homogeneity among populations. Most of the populations were in Hardy–Weinberg equilibrium. According to the fixation index, populations were panmictics. The partition of the F-statistic demonstrated that mating was at random, although there was a slight excess of heterozygotes, and there were no genetic differences among populations. The Shannon–Weaver index showed that more than 96% of the total genetic diversity for this locus is attributable to the differences within each herd. The diversity among herds within locations and among locations accounted for $< 4\%$ of the total variation. Phenograms obtained by the Nei procedure for genetic distances disclosed the separation of the herds into two groups, one with an average frequency of LGB B of 0·61 and the other with an LGB B frequency of 0·53. Locations defined two clusters, but the divergence among populations was not important. These results indicated that the Argentinian Holstein breed from Córdoba constitutes one population.

INTRODUCTION

The bovine protein β-lactoglobulin (LGB) represents 13% of the total protein in milk. Since Aschaffenburg & Drewry (1955) described the presence of different β-lactoglobulin variants (A, B, C and D), many studies have been carried out on the genetic polymorphism of this system in different bovine breeds. The allele LGB B predominates in most dairy breeds. Its frequency reaches *c*. 80% in Ayrshire, 70% in Guernsey and milk Shorthorn, 60% in Friesian, Brown Swiss, Jersey and Holstein, but in this latter breed, alleles LGB A and LGB B are, sometimes, in equal proportion (Li & Gaunt 1972; Ng-Kwai-Hang *et al*. 1984; Tejedor *et al*. 1986; Van Eenennaam & Medrano 1991). The variant LBG C is characteristic of Jersey cattle (McLean *et al*. 1984) and LGB D was found only in French breeds (Grosclaude 1988). The protein β-lactoglobulin was studied by Poli & Antonini (1991) in the Argentinian Holstein, which is the most important dairy breed in Argentina.

This protein has great economic value. Several studies from different breeds showed the association of LGB polymorphisms with variations in milk production, composition and processing properties of milk. Variants at the LGB locus had no significant effect on milk yield in the first lactation (McLean *et al*. 1984; Ng-Kwai-Hang *et al*. 1984) but Ng-Kwai-Hang *et al*. (1986) and Bovenhuis *et al*. (1992), from data collected over different lactations, demonstrated that higher milk yield was associated with LGB AA than with BB genotypes. Fat and casein concentration, casein number and total solids were significantly higher in milk from LGB BB than from AA genotypes. The total contents of proteins and whey proteins were higher from the LGB AA genotype than from the other genotypes. The LGB AB genotype was intermediate for those characteristics in some studies but in others showed the same behaviour as the LGB BB genotype (McLean *et al*. 1984; Ng-Kwai-Hang *et al*. 1984, 1986). Van Eenennaam & Medrano (1991) have found no effect of this system on milk yield traits. Also Lin *et al*. (1989) found that LGB genotypes had no significant effect on lifetime performance of dairy cattle. Milk with the LGB B variant is preferred for use in cheese manufacture because it has better coagulating properties than milk from LGB A cattle (Morini *et al*. 1982; Schaar *et al*. 1985; Marziali & Ng-Kwai-Hang 1986).

* To whom all correspondence should be addressed. Email: abonvillani@ayv.unrc.edu.ar

https://doi.org/10.1017/S0021859698005528 Published online by Cambridge University Press

Theoretical studies on population structure predict the occurrence of two categories regarding the distribution of variability. In species that are continuously distributed, intra-population variability is more prevalent than in those species which are made up of small isolated populations (Wright 1931; Kimura & Crow 1964; Nei *et al.* 1975). The objective of this study was to determine the population structure of LGB in the Argentinian Holstein breed from the Córdoba province, and to quantify the genetic variation employing several indices.

MATERIALS AND METHODS

Individual milk samples were obtained from 360 cows, 30 samples per herd were taken during 1995 from a total of 12 herds, in the four principal locations of cattle milk production: five at Coronel Moldes (CM); two at Villa María (VM); two at Devoto (DV) and three at La Carlota (LC).

The LGB variants were identified using alkaline polyacrylamide gel electrophoresis as described by Kiddy *et al.* (1972).

Allelic frequencies were determined by direct counting. The comparability of allelic and genotypic frequencies among herds and locations within the total population was evaluated by $m \times n$ contingency tables, where m (row) represents herds or locations and n (column) represents the allelic or genotypic frequencies. The G-test (Sokal & Rohlf 1979) was applied to find out if the populations were in Hardy–Weinberg equilibrium.

The unbiased fixation index (F) (Brown & Weir 1983) was calculated for each herd and location. The whole population was examined by the Wright F-statistic under different levels of hierarchical structure within the population for this locus, in agreement with Crow & Kimura (1970):

$$(1-F_{IT}) = (1-F_{IR})(1-F_{RS})(1-F_{ST})$$

where T is the total population studied, I is the individual within the herd, R is the herd and is the subpopulation of S, and S is the location of milk production, which is the subpopulation of T.

The Shannon–Weaver diversity index (I) (Lewontin 1972; Ibañez *et al.* 1994) was used to quantify the genetic diversity, within and among herds and locations. Electrophoretic allelic frequencies were analysed by the diversity index as follows:

$$I = -\sum_{i=1}^{n} p_i \log p_i \quad \text{being} \quad \sum_{i=1}^{n} p_i = 1$$

where p_i represents the relative allelic or genotypic frequencies at the LGB locus for each herd or location and n is the number of alleles or genotypes for this locus. The I index was calculated for the total population (I_T) as an average from allelic or genotypic frequencies over all the herds and it was apportioned considering the mean diversity (allelic or genotypic) from the herds (I_R) and the locations (I_S). The mean diversity was estimated as the arithmetic average of the I values over the twelve herds and for four locations.

Genetic distances among populations, 12 herds and four locations were estimated from their allelic distribution using the Nei distance index (Nei 1972) and were summarized in the form of a phenogram derived from the UPGMA (unweighted pair group method using arithmetic averages) clustering method (Sneath & Sokal 1973).

RESULTS AND DISCUSSION

Protein patterns

In accordance with the results obtained by Aschaffenburg & Drewry (1955) and Van Eenennaam & Medrano (1991), the electrophoresis of the monomeric protein β-lactoglobulin produced two bands corresponding with the action of the two autosomal codominant alleles from the LGB locus, allowing the identification of the two homozygotes and the heterozygote (Fig. 1).

Allelic and genotypic frequencies

Table 1 shows the observed genotypic frequencies and the estimated allelic frequencies for each herd and location, indicating that the allele frequencies of LGB A and LGB B were similar and therefore close to 0·5. The allelic and genotypic frequencies were also similar among herds and among locations so that the homogeneity χ^2 (Table 2) resulting from the $m \times n$ contingency tables did not produce any statistically significant difference. On the other hand, the G-test between the observed and expected genotype frequencies, assuming a Hardy–Weinberg proportion

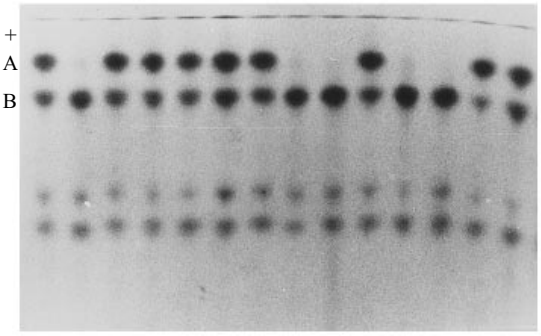


Fig. 1. Polyacrylamide alkaline gel electrophoresis of the β-lactoglobulin protein showing two bands corresponding to the action of the alleles A and B. Lanes that have two bands are heterozygotes AB and lanes that have only the slower band are homozygotes BB.

https://doi.org/10.1017/S0021859698005528 Published online by Cambridge University Press

Table 1. *Genotypic and allelic frequencies, equilibrium test (G) and fixation index (F) for each herd and location in Argentinian Holstein cattle*

Herds	AA	AB	BB	A	B	G	F
CM1	6	19	5	0·52	0·48	2·18	−0·25
CM2	4	13	13	0·35	0·65	0·07	0·21
CM3	8	12	10	0·47	0·53	1·16	0·21
CM4	8	9	13	0·42	0·58	4·46	0·40
CM5	3	18	9	0·60	0·40	1·93	−0·23
VM6	5	13	12	0·38	0·62	0·21	0·10
VM7	5	14	11	0·40	0·60	0·02	0·04
DV8	3	21	6	0·45	0·55	5·36	−0·40
DV9	5	19	6	0·48	0·52	2·18	−0·25
LC10	4	16	10	0·40	0·60	0·37	−0·09
LC11	8	11	11	0·45	0·55	2·03	−0·29
LC12	5	19	6	0·48	0·52	2·18	−0·25
Location	AA	AB	BB	A	B	G	F
CM	29	71	5	0·43	0·57	0·18	0·04
VM	10	27	23	0·39	0·61	0·19	0·06
DV	8	40	12	0·47	0·53	7·07	−0·33
LC	17	46	27	0·44	0·56	0·11	−0·03

CM: Coronel Moldes. VM: Villa María. DV: Devoto. LC: La Carlota.

Table 2. *Comparison of LGB allelic and genotypic frequencies by m × n contingency tables*

	Allelic			Genotypic		
	D.F.	χ^2	*P*	D.F.	χ^2	*P*
Among herds	11	6·5	0·83	22	27·4	0·20
Among locations	3	1·5	0·68	6	8·3	0·21

D.F. = degrees of freedom; χ^2 = Chi-squared from contingency tables; *P* = probability value.

(Table 1), did not produce evidence of significant differences except for the CM4 and DV8 herds and for the DV location.

The genotypic frequencies obtained showed a slight predominance of the heterozygous genotype and that the LGB B allele frequency of 0·57 was slightly higher than the LGB A allele, which is in agreement with Ng-Kwai-Hang *et al*. (1984), Tejedor *et al*. (1986), Poli & Antonini (1991) and Van Eenennaam & Medrano (1991). The majority of herds at different locations can be considered to be panmictic (random mating populations), where directional or stochastic factors are not operating. The departure from Hardy–Weinberg expectations of the genotypic frequencies corresponding to the CM4 herd is due to a higher proportion of homozygotes LGB BB and, in the case of the DV8 herd and DV location, is due to an excess of animals with a heterozygous genotype.

On the other hand, the great similarity among allelic and genotypic frequencies hinders the characterization or differentiation of herds and locations. The homogeneity among analysed samples allows us to consider the Argentinian Holstein breed from Córdoba as a single population.

Genetic structure

Genotypic structure

The fixation index (F) calculated for this locus, exhibited, in a great number of the herds, values close to zero and negative values in 50% of the herds and locations analysed (Table 1). From all herds, the highest positive coefficient was for CM4 (F = 0·40) and the highest negative was for DV8 (F = −0·40). Among all locations the most remarkable was DV (F = −0·33).

Total endogamy was $F_{IT} = -0{\cdot}046$, from the partition of the F-statistic indices considering the three hierarchical levels of structure: individual, herd and location, the coefficient values obtained were: $F_{IR} = -0{\cdot}055$; $F_{RS} = 0{\cdot}007$; $F_{ST} = 0{\cdot}002$.

The CM4 herd was the only one where a relatively high endogamy coefficient was observed. Herds and locations with a negative fixation index reveal the existence of an excess of heterozygotes. This higher than expected proportion of LGB AB animals was particularly observed in the DV8 herd and in the DV location. The G-test for equilibrium in those three cases was significant, this being the more probable cause of its deviation from Hardy–Weinberg expectations. Negative and positive values in equal proportion from the F index, like the analysed data, are characteristics of highly polymorphic and panmictic populations.

F_{IT} and F_{IR} values close to zero indicated that there were no effects of endogamy and that mating was at random, verifying that the genotypic frequencies from each herd, in general, agreed with those expected for panmictic populations. The negative values of both indices also indicated a slight excess of heterozygotes within each herd (R) and in the total population (T). Genetic drift could result from population subdivision of herds within locations (F_{RS}) or from locations within the total population (F_{ST}). Their low values indicate that endogamy is not operating and there are no genetic differences among herds within the location or among locations. The low F_{ST} values also suggest that gene flow occurs within the population.

The finding of negative fixation coefficients agrees with the data of Van Eenennaam & Medrano (1991), who reported for the first time a genotypic frequency deviating from equilibrium for this protein and in this breed, due to a greater number of heterozygotes than

https://doi.org/10.1017/S0021859698005528 Published online by Cambridge University Press

Table 3. *Apportionment of the Shannon–Weaver index calculated with the genotypic and allelic frequencies (values in %)*

	Allelic	Genotypic	
I_R/I_T	99·34	96·15	(Within R)
$(I_S–I_R)/I_T$	0·44	1·14	(Among R within S)
$(I_T–I_S)/I_T$	0·22	2·72	(Among S)

I_R, I_S, I_T = diversity indices from herds, locations and total population.

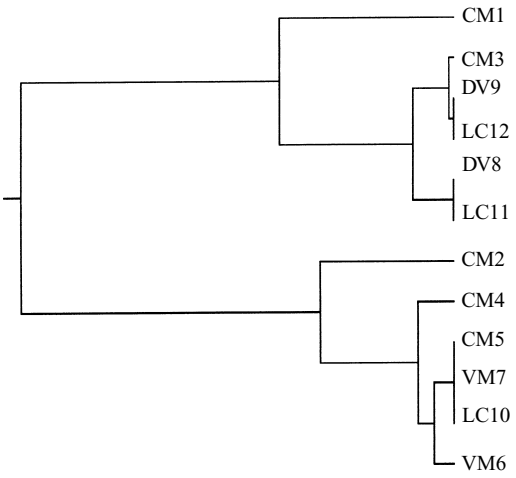


Fig. 2. Phenogram of the 12 herds based on the allelic frequencies at the LGB locus, employing the UPGMA clustering method.

expected. Li & Gaunt (1972) and Van Eenennaam & Medrano (1991) interpreted this observation as a selective advantage of one of the alleles. This could result from (*a*) selective advantage of the LGB AB genotype or (*b*) the indirect practice of negative assortative mating due to selection for metric characters according to industry requirements: high milk and fat yields. Several studies showed that these characters are associated with LGB AA and BB genotypes, respectively. Therefore selection would indirectly favour the persistence of the two variants A and B in the population, resulting in an increase of the heterozygote genotype frequency.

Allelic and genotypic diversity

The results obtained from the hierarchical analysis of the Shannon–Weaver index are shown in Table 3. Most of the total allelic diversity (99%) and the genotypic diversity (96%) were attributable to the genetic differences within each herd (I_R/I_T). The diversity among herds within location and among

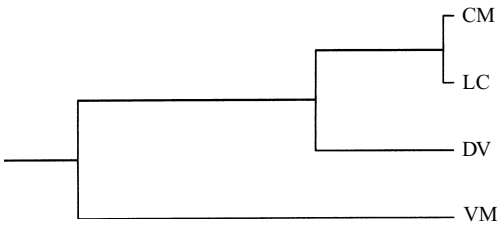


Fig. 3. Phenogram of the four locations based on the allelic frequencies at the LGB locus, employing the UPGMA clustering method.

locations, scarcely contributed (with $< 4\%$) to the total variation.

Genetic distances

The phenogram obtained according to the allelic frequencies of the LGB locus from the twelve herds allowed a separation into two groups of equal size, one with an average frequency of LGB B allele 0·61 and the other with LGB B 0·53 (Fig. 2). The distribution of herds within each group is at random, so there is no clustering by locations. The phenogram presenting the allelic frequencies from the four locations separated VM location from a cluster of DV, LC and CM, while DV diverges from LC and CM, both having small gene frequency differences (Fig. 3 and Table 1).

The failure to arrange herds in groups corroborated the inferences obtained with the Shannon–Weaver index, which indicated that among herds and location differences barely contributed to the total diversity. The grouping of CM, LC and DV locations is probably due to the fact that farms deal with the directives of the same milk industry firm. Although the divergence of those with the VM location is not very important, it is in agreement with the data and analyses from Table 1.

The results obtained from the analysis employing χ^2, *G*-test, endogamy coefficient, diversity index and distances index lead us to conclude that the greatest proportion of the genetic variability is contained within each herd, there were no genetic differences between herds or locations, and thus we can consider the Argentinian Holstein breed as a single population for the LGB locus in the Córdoba province.

On the other hand, in accordance with the allelic frequencies found in this study, it is feasible to select for one of the two alleles. Van Eenennaam & Medrano (1991) showed that over time and in different breeds, the frequencies of LGB and of other milk proteins did not change and have remained fixed during the last 30 years in spite of intensive selection for milk yield.

Even if the milk protein loci were related to desirable metric traits, the effects of assortative mating and truncation selection on changing allele frequencies of a single locus may be very small and other factors

https://doi.org/10.1017/S0021859698005528 Published online by Cambridge University Press

such as linkage disequilibrium may complicate the situation (Li & Gaunt 1972) or there could be a selection process acting on a locus closely linked to the LGB locus.

Nevertheless, the milk industry is interested in improving protein yield and quality. Thus, it should be possible that, in the near future, the application of more efficient selection methods and the coordination of farmers and industries to attain those goals would produce useful changes in the population.

REFERENCES

Aschaffenburg, R. & Drewry, J. (1955). Occurrence of different beta-lactoglobulins in cow's milk. *Nature* **176**, 218–219.

Bovenhuis, H., van Arendonk, J. A. M. & Korver, S. (1992). Associations between milk protein polymorphisms and milk production traits. *Journal of Dairy Science* **75**, 2549–2559.

Brown, A. H. & Weir, B. S. (1983). Isozymes in plant genetics and breeding. In *Plant Genetics and Breeding*, *Part A* (Eds S. D. Tanksley & T. J. Norton), pp. 219–240. Amsterdam: Elsevier Science Publishers.

Crow, F. J. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper & Row Publishers.

Grosclaude, F. (1988). Le polymorphisme génétique des principales lactoproteines bovines. *INRA Production Animales* **1**, 5–17.

Ibañez, M. A., Di Renzo, M. A. & Poverene, M. M. (1994). Isozymic polymorphisms in a diploid population of lovegrass (*Eragrostis curvula*). *Journal of Genetics and Breeding* **48**, 141–148.

Kiddy, C. A., Rollins, R. E. & Zikakis, J. P. (1972). Discontinuous polyacrylamide electrophoresis for β-lactoglobulin typing of cow's milk. *Journal of Dairy Science* **55**, 1506–1507.

Kimura, M. & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.

Lewontin, R. C. (1972). The apportionment of human diversity. *Evolutionary Biology* **6**, 381–398.

Li, H. F. & Gaunt, S. (1972). A study of genetic polymorphisms of milk β-lactoglobulin, αs1-casein, β-casein, and κ-casein in five dairy breeds. *Biochemical Genetics* **6**, 9–20.

Lin, C. Y., McAllister, A. J., Ng-Kwai-Hang, K. F., Hayes, J. F., Batra, T. R., Lee, A. J., Roy, G. L., Vesely, J. A., Wauthy, J. M. & Winter, K. A. (1989). Relationships of milk protein types to lifetime performance. *Journal of Dairy Science* **72**, 3085–3090.

Marziali, A. S. & Ng-Kwai-Hang, K. F. (1986). Effects of milk composition and genetic polymorphism on cheese composition. *Journal of Dairy Science* **69**, 2533–2542.

McLean, D. M., Graham, E. R. B., Ponzoni, R. W. & McKenzie, H. A. (1984). Effects of milk protein genetic variants on milk yield and composition. *Journal of Dairy Research* **51**, 531–546.

Morini, D., Castagnetti, G., Chiavaric, C., Losi, G., Davoli, R. & Bosi, P. (1982). Prova di caseificazione con latte caratterizzato dalle varianti A e B della Beta-lattoglobulina nella produzione del formaggio parmigiano-reggiano. *Scienza E Tecnica Lattiero Casearia* **33** (6), 475–492.

Nei, M. (1972). Genetic distance between populations. *American Naturalist* **106**, 283–292.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10.

Ng-Kwai-Hang, K. F., Hayes, J. F., Moxley, J. E. & Monardes, H. G. (1984). Association of genetic variants of casein and milk serum proteins with milk, fat, and protein production by dairy cattle. *Journal of Dairy Science* **67**, 835–840.

Ng-Kwai-Hang, K. F., Hayes, J. F., Moxley, J. E. & Monardes, H. G. (1986). Relationships between milk protein polymorphisms and major milk constituents in Holstein-Friesian cows. *Journal of Dairy Science* **69**, 22–26.

Poli, M. A. & Antonini, A. G. (1991). Genetic structure of milk proteins in Argentinian Holstein and Argentinian Creole cattle. *Hereditas* **115**, 177–182.

Schaar, J., Hansson, B. & Pettersson, H.-E. (1985). Effects of genetic variants of κ-Cn and β-lactoglobulin on cheesemaking. *Journal of Dairy Research* **52**, 429–437.

Sneath, P. H. A. & Sokal, R. R. (1973). *Numerical Taxonomy*. San Francisco: W. H. Freeman & Co.

Sokal, R. R. & Rohlf, F. J. (1979). *Biometría*. Spanish edition. Madrid: Ed. H. Blume.

Tejedor, T., Zarazaga, I. & Altarriba, J. (1986). Consideraciones sobre la mejora genética de las características tecnológicas de la leche en ganado Holstein-Friesian. *ITEA* **66**, 27–33.

Van Eenennaam, A. & Medrano, J. F. (1991). Milk protein polymorphisms in California dairy cattle. *Journal of Dairy Science* **74**, 1730–1742.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.

https://doi.org/10.1017/S0021859698005528 Published online by Cambridge University Press