*Genetics and Molecular Biology*, 29, 3, 491-495 (2006)

www.sbg.org.br

Research Article

# Analysis of the population structure of Uruguayan Creole cattle as inferred from milk major gene polymorphisms

Gonzalo Rincón, Eileen Armstrong and Alicia Postiglioni

*Laboratorio de Análisis Genéticos en Animales Domésticos, Facultad de Veterinaria, Area Genética, Montevideo, Uruguay.*

## Abstract

The ancestors of Uruguayan Creole cattle were introduced by the Spanish conquerors in the XVII century, following which the population grew extensively and became semi-feral before the introduction of selected breeds. Today the Uruguayan Creole cattle genetic reserve consists of 575 animals. We used the tetra primer amplification refractory mutation system polymerase chain reaction (ARMS-PCR) to analyze the κ-*casein*, β-*casein*, α*S1-casein* and α-*lactoalbumin* gene polymorphisms and restriction fragment length polymorphism PCR (RFLP-PCR) for the β-*lactoglobulin* and the *acylCoA:diacyl glycerol acyltransferase 1* (*DGAT1*) genes. The κ-*casein* and β-*lactoglobulin* genes presented very similar A and B allele frequencies, while the α*s1-casein* and α-*lactoalbumin* gene B alleles showed much higher frequencies than the corresponding A alleles. The β-*casein* B allele was not found in the population sampled. There was a very high frequency of the *DGAT1* gene A allele which is associated with low milk fat content and high milk yield. All loci were in Hardy-Weinberg equilibrium and the level of heterozygosity agreed with the high genetic diversity observed in a previous analysis of this population. Preservation of the allelic richness observed in the Uruguayan Creole cattle should be considered for future dairy management and livestock genetic improvement. The results also emphasize the value of the tetra primers ARMS-PCR technique as a rapid, easy and economical way of genotyping cattle breeds for milk gene single nucleotide polymorphisms.

*Key words:* DGAT1 gene, milk protein, SNPs, Uruguayan Creole cattle.

Received: June 1, 2005; Accepted: December 14, 2005.

## Introduction

Cattle were first introduced to Uruguay In the XVII century by a Spanish conqueror named Hernando Arias de Saavedra from the Iberian Peninsula and later were also brought from the Jesuit missions located in a region called `Alto Uruguay', these two foundation stocks resulting in the establishment of the Uruguayan Creole cattle (UCC) population (Wilkins *et al*., 1989; Primo, 1992). During the XIX century, several commercial European cattle breeds were brought to America, leading to genetic introgression that greatly reduced the Creole population.

In 1930, a group of 35 Creole bulls, cows and calves were brought from inhabited areas of Treinta y Tres and Maldonado departments and established what is today the Uruguayan Creole cattle reserve (Arredondo, 1958; Postiglioni *et al*., 2002). Characterization of this population has been performed with different markers. As a first step, basic morphological (such as coat color, horn shape) and cytogenetic markers were studied. The presence of the robertsonian translocation 1;29 and absence of the acrocentric Y chromosome observed in *Bos taurus indicus* together with the analysis of various molecular markers supported the Iberian origin of Uruguayan Creole cattle as typical *Bos taurus taurus* (Rodríguez *et al*., 2001; Postiglioni *et al*., 2002).

Random amplified polymorphic DNA (RAPD) markers have been used to analyze random genome regions of Uruguayan Creole cattle and have revealed differences in band-sharing frequencies and specific banding patterns in Creole cattle as compared to selected European breeds such as Holstein-Friesian and Hereford (Rincón *et al*., 2000; Postiglioni *et al*., 2002). Later research detected 22 different alleles in the second exon of the *DRB3* gene, a polymorphic region related to the immune response of the BoLa gene (Postiglioni *et al*., 2002; Kelly *et al*., 2003). The analysis of 17 microsatellites showed an expected heterozygosity per locus of between 0.46 and 0.80, except for the *HEL13* locus (He=0.29). The high expected mean heterozygosity (0.62) and the low mean $F_{IS}$ index (0.098) suggest

Send correspondence to Alicia Postiglioni. Laboratorio de Análisis Genéticos en Animales Domésticos, Facultad de Veterinaria, Area Genética, Avda. A. Lasplaces 1550, 11600 Montevideo, Uruguay. E-mail: alipos@adinet.com.uy.

that the population is able to self-sustain on the long run (Armstrong, 2004).

It is now accepted that the potential value of local livestock breeds must be analyzed and conserved so that they can become a self-sustainable resource. The importance for conservation of specific alleles or genotypes and allelic richness has been also advocated (Gandini and Villa, 2003).

Milk protein genetic variability in cattle has been extensively studied at both the DNA and protein levels for evolutionary and biodiversity analyses (Caroli *et al*. 2004). Recent advances in single nucleotide polymorphism (SNP) detection allow genotyping without using restriction enzymes, a simpler and more economical way of detecting polymorphisms. One of these techniques, the tetra-primer amplification refractory mutation system polymerase chain reaction (ARMS-PCR) has been optimized to genotype SNPs in the bovine κ-*casein*, β-*casein*, α*S1-casein* and α-*lactoalbumin* genes (Rincón and Medrano, 2003). This technique uses a set of four primers in a single PCR reaction tube without the need for post-PCR manipulations to accomplish an efficient classification.

The *acylCoA:diacyl glycerol acyltransferase 1* (*DGAT1*) gene encodes the enzyme catalyzing the final step of triglyceride synthesis and has become a functional candidate gene for milk fat content after evidence that an increase in this trait in different breeds is strongly associated with a lysine at position 232 of the DGAT1 protein while an alanine at this position is associated with decreased milk fat content (Winter *et al*., 2002; Kuhn *et al*., 2004). The *DGAT1* gene has also been related to intramuscular fat deposition in cattle as it is mapped within the region of the marbling quantitative trait loci (QTL) (Thaller *et al*., 2003).

This paper outlines the results of a population structure analysis of Uruguayan Creole cattle using gene SNPs related to milk production traits. The tetra primer ARMS-PCR technique was used for the κ-*casein*, β-*casein*, α*S1-casein* and α-*lactoalbumin* genes while the restriction fragment length polymorphism PCR (RFLP-PCR) method was used for the β-*lactoglobulin* and *DGAT1* genes. This study provides the first characterization of the different alleles of the αS1-*casein*, β-*casein* and *DGAT1* genes in Uruguayan Creole cattle.

## Material and Methods

The population of Uruguayan Creole cattle of San Miguel National Park consists of 23 bulls, approximately 447 cows and 105 calves of both sexes. The present study was carried out on a total of 115 genomic DNA samples collected at random.

The tetra-primer ARMS-PCR procedure (Rincón and Medrano, 2003) was performed for the milk protein genes κ-*casein* (κ-*CN*), β-*casein* (β-*CN*), α*S1-casein* (α$s_1$-*CN*) and α-*lactoalbumin* (α-*LA*). The outer primers used for the κ-*casein*, α*S1-casein* and α-*lactoalbumin* were the same as those used in the RFLP-PCR technique, but the inner primers were designed by introducing a second deliberate mismatch at position -2 from the 3' end. The polymorphism K232A (AAG ⊗ GCG) in the *DGAT1* gene was analyzed using the RFLP-PCR technique (Winter *et al*, 2002). The β-*lactoglobulin* (β-*LG*) SNP was also studied using the RFLP-PCR method (Medrano and Aguilar-Cordoba, 1990).

For the statistical and genetic structure analysis we calculated the allele frequencies at the six loci analyzed, tested for Hardy-Weinberg equilibrium and estimated the observed (Ho), expected (He) and expected unbiased ($He_{unbiased}$) heterozygosity using the POPGENE32 program (Yeh *et al*., 1999).

Wright's fixation index ($F_{IS}$) was used as a measure of heterozygote deficiency or excess. Together with heterozygosity and the estimations of Hardy-Weinberg equilibrium, these indexes allowed further comprehension of the reproductive structure of the population analyzed.

## Results

Table 1 shows the allele frequencies detected at the six loci using the tetra-primers- ARMS-PCR and RFLP-PCR techniques. With the exception of the β-*CN* gene, two alleles were found in all the loci studied.

The κ-*CN* and β-*LG* genes presented very similar frequencies for both alleles (A and B) but the α$s_1$-*CN* gene B allele showed a much higher frequency than the C allele. For the α-*LA* gene, the B allele was present at a higher frequency than the allele A. The technique used for the β-*CN* gene only allowed us to differentiate between B and non-B alleles, therefore we assumed that A allele corresponded to any other allele different from the B allele.

In the case of the two *DGAT1* alleles we found that the A allele, corresponding to an alanine mutation associated with low milk fat content and high milk yield, was

**Table 1** - Allele frequencies and Hardy-Weinberg equilibrium (χ2 and p) for milk protein genes (κ-*casein*, β-*lactoglobulin*, α$S_1$-*casein*, α-*lactoalbumin*) and the *acyl-CoA:diacylglycerol acyltransferase* (*DGAT1*) gene.

| Allele | Gene name and allele frequency | | | | |
|---|---|---|---|---|---|
| | κ-*casein* | β-*lactoglobulin* | α$S_1$-*casein* | α-*lactoalbumin* | *DGAT1* |
| A | 0.5000 | 0.4938 | - | 0.2927 | 0.8864 |
| B | 0.5000 | 0.5062 | 0.8654 | 0.7073 | - |
| C | - | - | 0.1346 | - | - |
| K | - | - | - | - | 0.1136 |
| χ2 | 0.0680 | 0.0280 | 0.5300 | 0.2020 | 0.6443 |
| p | 0.7934 | 0.8650 | 0.4664 | 0.6531 | 0.4221 |

present at a very high frequency, while the K allele, considered the ancestral, showed a much lower frequency (Table 1).

The likelihood ratio test revealed that all the loci analyzed showed no significant departure from the expected Hardy-Weinberg equilibrium ($p > 0.05$; Table 1).

Observed, expected and mean values of heterozygosity and $F_{IS}$ statistics are shown in Table 2. The loci that presented both alleles at frequencies of around 0.50 had higher heterozygosities (κ-*CN* and β-*LG*), while loci showing one allele at a much higher frequency than the other (α*S1-CN* and *DGAT1*) showed lower values. The α-*LA* locus presented intermediate values.

Wright's fixation index ($F_{IS}$), a measure of the inbreeding coefficient, was in agreement with the other results, being low when genetic diversity was high (as in κ-*CN* and β-*LG*, $F_{IS} < 0.05$) and medium or high when the heterozygosity was low (α*S1-CN* and *DGAT1*, $F_{IS} > 0.10$). In these two cases a bias towards heterozygote excess was inferred from the negative $F_{IS}$ values that was confirmed by the differences between observed and expected heterozygosity. The observed high $F_{IS}$ values were not associated with Hardy-Weinberg disequilibrium.

## Discussion

Our study used milk major gene polymorphisms to elucidate the genetic structure of Uruguayan Creole cattle. Tetra-primer ARMS-PCR was used to successfully identify κ-*CN*, β-*CN*, $\alpha_{s1}$-*CN* and α-*LA* allele variants, while the RFLP-PCR method proved suitable for efficiently determining specific mutations in the β-*LG* and *DGAT1* genes.

Analyses performed on Holstein-Friesian cattle demonstrated that the B alleles for most of the proteins of the casein cluster as well as for β-*LG* are related to high cheese quality since they increase the rate of curd formation, rennet clotting time and coagulum strength (Van Eenennam and Medrano, 1990). We found that Uruguayan Creole cattle showed similar or even higher frequencies for the B allele for these protein genes. In the case of the β-*CN* gene the fact that we were unable to find any B allele implied that the allele is not present in the population or is present at a very low frequency as has been reported for many other cattle breeds (Jann *et al*., 2004).

Allelic variants of the κ-*CN*, $\alpha_{s1}$-*CN* and β-*LG* genes have also been analyzed in Argentinean and Bolivian Creole cattle (Lirón *et al*., 2002) and show similar frequencies to that observed in the Uruguayan Creole cattle. However, the Uruguayan Creole cattle allele distribution is even more similar to that seen in semi-wild Argentinean cattle (the Patagonian Creole) and to a Bolivian population (Saavedreño Creole) bred for dairy and beef production. Also, the expected heterozygosity of the κ-*CN* (0.5028) and β-*LG* (0.5031) genes in Uruguayan Creole cattle is higher than that occurring in Argentinean and Bolivian Creole herds (Lirón *et al*., 2002).

The $\alpha_{s1}$-*CN* polymorphism showed a high B allele frequency (0.8) and a low heterozygosity value (He = 0.2376) similar to that observed in Creole populations as well as in North American Holstein-Friesian cattle (Van Eenennaam and Medrano, 1990; Lirón *et al*., 2002).

The α-*LA* gene is related to the biosynthesis of lactose in the mammary glands and thus regulates the volume of milk, the effect of two allelic variants in the regulatory region of this gene having been studied by Bleck and Bremel (1993) who showed a positive correlation between the A allele and high milk yield, low milk fat and protein, while the BB genotype had a higher percentage of protein and fat. The A allele is not found in some beef and dairy breeds, but is thought to be involved in the expression of α-*LA* and thus in modulating milk production. The tetra primer ARMS-PCR genotyping technique detected both alleles and a medium level of heterozygosity (He = 0.4192) in the Uruguayan Creole cattle. This gene should be considered as an important target to preserve in this population for future dairy management.

In general, the level of heterozygosity found by us is in agreement with the high genetic diversity that has been previously observed in this population using other molecular markers (Rincón *et al*., 2000; Armstrong, 2004).

The *DGAT1* gene also presented both alleles described in the literature (Grisart *et al*., 2004; Winter *et al*., 2002). In cattle, allele frequency distribution show a tendency towards high frequencies of the *DGAT1* A allele in *Bos taurus* breeds and the same effect for the *DGAT1* K in *Bos indicus*. In Uruguayan Creole cattle, the *DGAT1* allele A (corresponding to the alanine mutation) was the more frequent (0.8864) than the ancestral K allele (0.1136) form. Considering the history and development of the Uruguayan Creole cattle population, this finding could be explained by genetic drift effects acting upon the population established around 1930 (Arredondo, 1958) as well as by the the founder effect.

**Table 2** - Heterozygosity (observed, Ho; expected, He; and unbiased, $He_{unbiased}$) and Wright's fixation index ($F_{IS}$) for the milk protein genes (κ-*casein*, β-*lactoglobulin*, $\alpha S_1$-*casein*, α-*lactoalbumin*) and the *acyl-CoA:diacylglycerol acyltransferase* (*DGAT1*) gene.

| Locus | Sample size (N)* | Heterozygosity | | $He_{unbiased}$ | $F_{IS}$ |
|---|---|---|---|---|---|
| | | Ho | He | | |
| κ-casein | 182 | 0.5165 | 0.5028 | 0.5000 | -0.0330 |
| β-lactoglobulin | 160 | 0.5125 | 0.5031 | 0.4999 | -0.0252 |
| $\alpha S_1$-casein | 52 | 0.2692 | 0.2376 | 0.2330 | -0.1556 |
| α-lactoalbumin | 82 | 0.3902 | 0.4192 | 0.4140 | 0.0575 |
| DGAT1 | 88 | 0.2273 | 0.2038 | 0.2014 | -0.1282 |
| Mean | 109 | 0.3193 | 0.3110 | 0.3081 | - |

*number of gene copies analyzed.

The *DGAT1* gene is considered to be a quantitative trait locus (QTL) for milk yield and composition, with the A allele being related to low milk fat and high milk protein content, high milk-yield (Winter *et al*., 2002; Spelman *et al*.; 2002; Grisart *et al*., 2004) and low intramuscular fat (marbling) (Thaller *et al*, 2003). We found that Uruguayan Creole cattle showed similar *DGAT1* allele frequencies when compared to a beef breed such as Aberdeen Angus (A allele = 0.87; K allele = 0.13) and the old Spanish breed Toro de Lidia (A = 0.79; K = 0.21) considered closely related to American Creole cattle (Kaupe *et al*. 2003). Breeds selected for milk production show more variation in *DGAT1* allele frequencies, some being near to 0.50 for each allele (*e.g.* German Holstein, A allele = 0.58; K allele :0.42) or higher for the K allele (*e.g.* New Zealand Holstein-Friesian, A allele = 0.40; K allele 0.60 and Jersey, A allele = 0.31; K allele = 0.69), which is most likely the result of artificial selection for increasing milk yield and/or high milk fat content (Kaupe *et al*., 2003; Spelman *et al*., 2002).

Low milk fat content is related to an increase in follicular activity of the ovaries, which results in a higher fertility of the female (Lucy *et al*., 1992). As the production of low fat milk also results in lower energy expenditure these two factors combined may give a natural selective advantage for cattle which carry the A allele (Kaupe *et al*., 2004). This is an even more interesting observation when applied to a semi-wild unselected population of cattle such the Uruguayan Creole cattle herd because these animals have lived under natural conditions with almost no management for 400 years (Arredondo, 1958; Postiglioni *et al*., 2002; Rincón *et al*., 2000; Armstrong, 2004).

All loci were in Hardy-Weinberg equilibrium which is in agreement with previous protein loci data obtained from this cattle population (Postiglioni *et al*., 2002). However, medium or high $F_{IS}$ values in the two loci that show very low frequencies of one of the detected alleles (*DGAT1* and $\alpha S_1$*-CN*) should not be overlooked. These two loci also showed lower gene diversity (measured as He), due to the large difference between allele frequencies. The heterozygosity level of the *DGAT1* found in the Creole cattle was similar to that found in other breeds that exhibit similar allele frequencies (*e.g.* Toro de Lidia, He = 0.33; Aberdeen Angus, He = 0.21) (Kaupe *et al*.; 2003).

The results presented in this paper show the value of the tetra primers ARMS-PCR technique as a quick, easy and economical way of genotyping cattle breeds for milk related gene SNPs. Another important finding of this work is the allelic richness and high level of heterozygosity of most molecular markers studied in the Uruguayan Creole cattle, which shows that this reserve is a precious source of genetic variation that should be maintained or used in livestock genetic improvement.

## Acknowledgments

The authors wish to thank the following: Mrs. Iris Hernández for technical assistance; DVM Marcela Silveira and personnel of the Servicio de Parques del Ejército (SEPAE) for their help in collecting samples; and personnel at the Animal Genomics Laboratory at UCDavis, University of California for the genotyping techniques. This research was supported by grants from Programa de Desarrollo de las Ciencias Básicas (PEDECIBA), Universidad de la República (CSIC) and Facultad de Veterinaria (CIDEC).

## References

Armstrong E (2004) Análisis de la diversidad genética del bovino Criollo Uruguayo mediante microsatélites. MSc Thesis, Universidad de la República, Uruguay.

Arredondo H (1958) Santa Teresa y San Miguel. La Restauración de las Fortalezas. La Formación de sus Parques. Imprenta "El Siglo Ilustrado", Montevideo.

Bleck GT and Bremel RD (1993) Correlation of the α-lactoalbumin 5' flanking region. J Dairy Sci 76:2292-2298.

Caroli A, Chessa S, Bolla P, Budelli E and Gandini GC (2004) Genetic structure of milk protein polymorphisms and effects on milk production traits in a local dairy cattle. J Anim Breed Genet 121:119-127.

Gandini GC and Villa E (2003) Analysis of the cultural value of local livestock breeds: A methodology. J Anim Breed Genet 120:1-11.

Grisart B, Farnir F, Karim L, Cambisano N, Kim J, Kvasz A, Mni M, Simori P, Frere J, Coppieters W and Georges M (2004) Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc Natl Acad Sci USA 101:2308-2403.

Jann OC, Ibeagha-Awemu EM, Ozbeyaz C, Zaragoza P, Williams JL, Ajmone-Marsan P, Lenstra JA, Moazami-Goudarzi K and Erhardt G (2004) Geographic distribution of haplotype diversity at the bovine casein locus. Genet Sel Evol 36:243-257.

Kaupe B, Winter A, Fries R and Erhardt G (2004) DGAT1 polymorphism in *Bos Indicus* and *Bos taurus* cattle breeds. J Dairy Res 71:182-187.

Kelly L, Nicolini P, D'Angelo M, Nimo A, Rincón G, Piaggio J and Postilglioni A (2003) Polimorfismo del gen DRB3.2 en bovinos Criollos del Uruguay. Archivos de Zootecnia 52:77-80.

Kuhn CH, Thaller G, Winter A, Bininda-Emonds O, Kaupe B, Erhardt G, Bennewitz J, Schwerin M and Fries R (2004) Evidence for multiple alleles at the DGAT1 locus better explains a quantitative trait locus with major effect on milk fat content in cattle. Genetics 167:1873-1881.

Lirón JP, Ripoli MV, De Luca JC, Peral-García P and Giovambattista G (2002) Analysis of genetic diversity and population structure in Argentine and Bolivian Creole cattle using five loci related to milk production. Genet Mol Biol 25:413-419.

Lucy MC, Savio JD, Badinga L, De La Sota RL and Thatchers WW (1992) Factors that affect ovarian follicular dynamics in cattle. J Anim Sci 70:3615-3626.

Medrano JF and Aguilar-Córdova E (1990*)* Polymerase chain reaction amplification of bovine β-lactoglobulin genomic sequences and identification of genetic variants by RFLP analysis. Anim Biotechnol 1:73-77.

Postiglioni A, Rincón G, Kelly L, Llambí S, Fernández G, D'Angelo M, Gagliardi R, Trujillo J, de Bethencourt M, Guevara K, Castellano A and Arruga MV (2002) Biodiversidad genética en bovinos Criollos del Uruguay. Análisis con marcadores moleculares. Archivos de Zootecnia 51:1-8.

Primo AT (1992) El ganado bovino ibérico en las Américas: 500 años después. Archivos de Zootecnia 41(extra):421-432.

Rincón G, D'Angelo M, Gagliardi R, Kelly L, Llambí S and Postiglioni A (2000) Genomic polymorphism in Uruguayan Creole cattle using RAPD and microsatellite markers. Res Vet Sci 69:171-174.

Rincón G and Medrano JF (2003) Single nucleotide polymorphism genotyping of bovine milk protein genes using the tetra-primer ARMS-PCR. J Anim Breed Genet 120:331-337.

Rodríguez M, Fernandez G, Silveira C and Delgado JV (2001) Estudio étnico de los bovinos Criollos del Uruguay. I. Análisis biométrico. Archivos de Zootecnia 50:113-118.

Spelman RJ, Ford CA, McElhinney P, Gregory GC and Snell RG (2002) Characterization of the DGAT1 gene in the New Zealand dairy population. J Dairy Sci 85:3514-3517.

Thaller G, Kuhn C, Winter A, Ewlad G, Bellmann O, Wegner J, Zuhlke H and Fries R (2003) DGAT1, a new positional and functional candidate gene for Intramuscular fat deposition in cattle. Anim Genet 34:354-357.

Van Eenennaam A and Medrano JF (1990) Milk protein polymorphisms in California Dairy cattle. J Dairy Sci 74:1730-1742.

Wilkins JV, Martinez L and Rojas F (1989) El ganado vacuno Criollo. Diálogo XXXVI. Conservación y mejoramiento del ganado bovino Criollo IICA. PROCISUR:69-82.

Winter A, Kramer W, Werner F, Kollers S, Kata S, Durstewitz G, Buitkamp J, Womack W, Thaller G and Fries R (2002) Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proc Natl Acad Sci USA 99:9300-9305.

Yeh FC, Boyle T, Rongcai Y, Ye Z and Mao Xiyan J (1999) PopGene v.1.31. Microsoft window-based freeware for population Genetic Analysis. Department of Renewable Resources, University of Alberta, Canada.

*Associate Editor: Pedro Franklin Barbosa*