ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12529

Genetic diversity and patterns of population structure in Creole goats from the Americas

C. Ginja*, L. T. Gama†, A. Martínez‡, N. Sevane§, I. Martin-Burriel¶, M. R. Lanari, M. A. Revidatti††, J. A. Aranguren-Méndez‡‡, D. O. Bedotti§§, M. N. Ribeiro¶¶, P. Sponenberg***, E. L. Aguirre†††§§§§, L. A. Alvarez-Franco‡‡‡, M. P. C. Menezes§§§, E. Chacón¶¶¶, A. Galarza****, N. Gómez-Urviola††††, O. R. Martínez-López‡‡‡‡, E. C. Pimenta-Filho§§§, L. L. da Rocha¶¶, A. Stemmer****, V. Landi‡ and J. V. Delgado-Bermejo‡**

*CIBIO-InBIO – Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas n. 7, 4485-661 Vairão, Portugal. †CIISA, Faculdade de Medicina Veterinária, Universidade de Lisboa, Lisboa, Portugal. ‡Departamento de Genética, Universidad de Córdoba, Campus de Excelencia Internacional Agroalimentario ceiA3, Córdoba, Spain. §Departamento de Producción Animal, Universidad Complutense de Madrid, Madrid, Spain. ¶Martin-Burriel, Laboratorio de Genética Bioquímica, Facultad de Veterinaria, Universidad de Zaragoza, Zaragoza, Spain. **Area de Producción Animal, Instituto Nacional de Tecnología Agropecuaria EEA, Bariloche, Argentina. ††Facultad de Ciencias Veterinarias, Universidad Nacional del Nordeste, Corrientes, Argentina. ‡‡Facultad de Ciencias Veterinarias, Universidad de Zulia, Maracaibo-Zulia, Venezuela. §§Instituto Nacional de Tecnología Agropecuaria EEA Anguil "Ing. Agr. Guillermo Covas", Bariloche, Argentina. ¶¶Departamento de Zootecnia, Universidade Federal Rural de Pernambuco, Recife, PE, Brazil. ***Virginia-Maryland Regional College of Veterinary Medicine, Virginia Tech, Blacksburg, VA, USA. †††Universidad Nacional de Loja, Loja, Ecuador. ‡‡‡Universidad Nacional de Colombia, Sede Palmira, Colombia. §§§Universidade Federal da Paraíba, Areia, PB, Brazil. ¶¶¶Universidad Técnica de Cotopaxi, La Maná, Ecuador. ****Universidad Mayor de San Simón, Cochabamba, Bolivia. ††††Universidad Nacional Micaela Bastidas de Apurímac, Abancay, Perú. ‡‡‡‡Centro Multidisciplinario de Investigaciones Tecnológicas, Dirección General de Investigación Científica y Tecnológica, Universidad Nacional de Asunción, San Lorenzo, Paraguay. §§§§Grupo de Melhoramento Animal e Biotecnologias GMAB-FZEA-USP, Brazil.

Summary

Biodiversity studies are more efficient when large numbers of breeds belonging to several countries are involved, as they allow for an in-depth analysis of the within- and between-breed components of genetic diversity. A set of 21 microsatellites was used to investigate the genetic composition of 24 Creole goat breeds (910 animals) from 10 countries to estimate levels of genetic variability, infer population structure and understand genetic relationships among populations across the American continent. Three commercial transboundary breeds were included in the analyses to investigate admixture with Creole goats. Overall, the genetic diversity of Creole populations (mean number of alleles = 5.82 ± 1.14, observed heterozygosity = 0.585 ± 0.074) was moderate and slightly lower than what was detected in other studies with breeds from other regions. The Bayesian clustering analysis without prior information on source populations identified 22 breed clusters. Three groups comprised more than one population, namely from Brazil (Azul and Graúna; Moxotó and Repartida) and Argentina (Long and shorthair Chilluda, Pampeana Colorada and Angora-type goat). Substructure was found in Criolla Paraguaya. When prior information on sample origin was considered, 92% of the individuals were assigned to the source population (threshold $q \geq 0.700$). Creole breeds are well-differentiated entities (mean coefficient of genetic differentiation = 0.111 ± 0.048, with the exception of isolated island populations). Dilution from admixture with commercial transboundary breeds appears to be negligible. Significant levels of inbreeding were detected (inbreeding coefficient > 0 in most Creole goat populations, $P < 0.05$). Our results provide a broad perspective on the extant genetic diversity of Creole goats, however further studies are needed to understand whether the observed geographical patterns of population structure may reflect the mode of goat colonization in the Americas.

Keywords admixture, biodiversity, domestic goats, genetic structure, short tandem repeats

Address for correspondence

C. Ginja, CIBIO-InBIO – Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas n. 7, 4485-661 Vairão, Portugal.
E-mail: catarinaginja@cibio.up.pt

Accepted for publication 02 November 2016

© 2017 Stichting International Foundation for Animal Genetics

Introduction

Goats arrived in America during the second Columbus trip in 1493 (Rodero *et al.* 1992). Due to the adaptability of this species, they were promptly spread throughout the entire continent and easily integrated into the farming activities of native civilizations, which until then were rearing mostly South American camelids. Natives and colonizers jointly contributed to the spread of goats throughout the continent, and there are now approximately 36 million goats in Latin America and the Caribbean and about 3 million in North America, representing about 3.5% of the world census (FAO 2007). Over time, several new goat populations were established in very distinct ecosystems on the American continents. This probably represented the major evolutionary event for the species subsequent to their domestication in the Middle East, approximately 10 000 years ago (Naderi *et al.* 2008). Goats in the Americas belong to mostly native breeds broadly called Creole, presumably with origins in the Iberian goats that arrived in the continent since the 15th century, during the period of European conquest and colonization (Rodero *et al.* 1992).

In spite of the social and economic relevance of American goats, these animals are poorly characterized. According to FAO, there are only about 28 breeds recognized in the region (Boettcher *et al.* 2014). The vast majority of native American goats are referred to as Creole, a reservoir of genetic diversity that has not been characterized. Creole goats are important resources available for endogenous rural development, especially in marginal areas and under a climatic change scenario, but they may be threatened by crossbreeding with and/or replacement by commercial transboundary breeds.

Molecular biology provides tools for the study of within- and between-population genetic diversity. As with other livestock species, short tandem repeats (i.e. microsatellites) have thus far been the markers of choice for genetic diversity studies (Groeneveld *et al.* 2010; Lenstra *et al.* 2012). Several results concerning the survey of goat genetic diversity based on microsatellites have been published, including those surveying breeds from Asia (Li *et al.* 2002; Fatima *et al.* 2008; Ağaoğlu & Ertuğrul 2012; Sulabda *et al.* 2012), Europe (Iamartino *et al.* 2005; Serrano *et al.* 2009; Bruno-de-Sousa *et al.* 2011; Martinez *et al.* 2015) and Africa (Traoré *et al.* 2009; Missohou *et al.* 2011), and also breeds from Latin American countries such as Brazil (Menezes *et al.* 2006), Colombia (Calvo *et al.* 2012) and Cuba (Chacón *et al.* 2012). The recent development of a SNP array for goats (Tosser-Klopp *et al.* 2014) will probably lead to the replacement of microsatellites by these markers in the future, as discussed in a comprehensive review of genetic diversity studies in goats (Ajmone-Marsan *et al.* 2014). The use of SNP arrays and genome-wide next-generation sequencing methodologies are becoming more common. Nonetheless, microsatellites seem to be adequate for estimating genetic diversity and depicting breed relationships (Bruford *et al.* 2015), and more so when a large number of animals and populations are genotyped to allow for more comprehensive inferences regarding their genetic composition.

Few studies with microsatellites concerning the genetic relationship among goat breeds from more than one country and across different climatic regions have been published thus far (Oliveira *et al.* 2010; Ribeiro *et al.* 2012). Nevertheless, biodiversity studies are more efficient when a large number of breeds belonging to several countries is involved, as this allows for an in-depth analysis of the within- and between-breed components of genetic diversity and inference on phylogenetic relationships. One example is the study by Cañon *et al.* (2006), who carried out a broad microsatellite-based analysis of European and Middle Eastern goat breeds within the ECONOGENE consortium and detected an East–West cline in genetic diversity along the Mediterranean.

This study was carried-out in the framework of the CONBIAND Network (http://www.uco.es/conbiand/Bienvenida.html), which is an international scientific association dedicated to the characterization and conservation of zoogenetic resources from Iberoamerica. Within this network, several studies have been conducted to assess the genetic diversity of farm animals of the American continent, including cattle (Ginja *et al.* 2010, 2013; Delgado *et al.* 2012) and pigs (Revidatti *et al.* 2014). In the context of the BioGoat Consortium (http://biogoat.jimdo.com/), we designed a sampling strategy to capture historical signatures of goat introductions into the Americas, including populations from the Caribbean Islands (Cuba), which represents the point of arrival of European settlers in the 15th century, and North America (United States) as well as South America (Colombia, Venezuela, Ecuador, Peru, Bolivia, Brazil, Paraguay, Argentina), following the routes of goat dissemination throughout the Americas (Primo 1992; Rodero *et al.* 1992; Villalobos Cortés *et al.* 2011). As a first approach, we used microsatellites to investigate the genetic composition of 24 goat breeds from 10 American countries, with the aims to: (i) characterize their levels of genetic variability and population structure and (ii) understand the patterns of genetic relationships among goat populations across the American continent.

Material and methods

Sampling and DNA extraction

Blood or hair samples were collected from a total of 910 animals of 24 Creole goat populations from the Americas (breed acronym, sample size): Argentina – Angora-type goat (TAN, 14), Chilluda longhair (CHL, 12), Chilluda shorthair (CHC, 11), Criolla del Nordeste (NEA, 39), Neuquina (NUQ, 50) and Pampeana Colorada (PCA, 14);

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

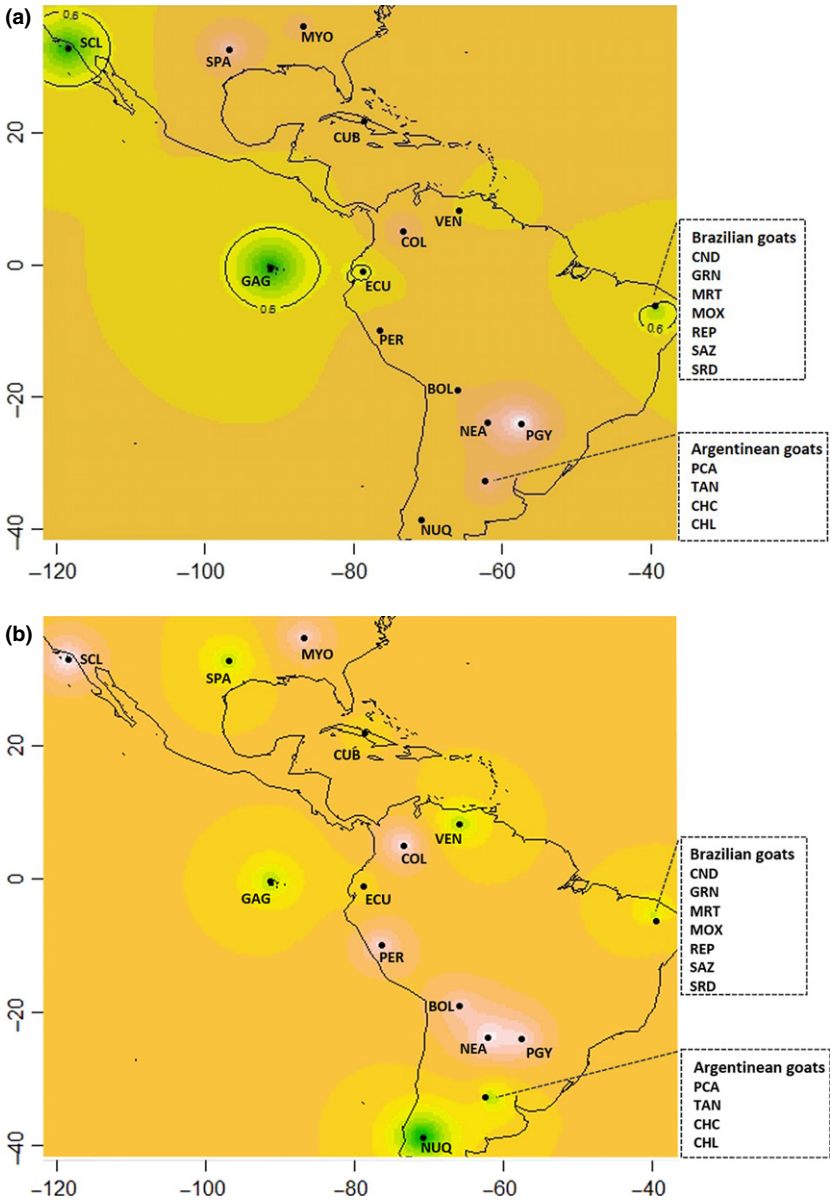


Figure 1 Contour graphics of geographical coordinates of Creole goat populations in relation to the estimated (a) H_e and (b) F_{is} values. Green areas correspond to the lower H_e and F_{is} values; pink areas correspond to the higher H_e and F_{is} values. Abbreviation of breed names and sample sizes are as follows: Argentina – Angora-type Goat (TAN, 14), Chilluda longhair (CHL, 12), Chilluda shorthair (CHC, 11), Criolla del Nordeste (NEA, 39), Neuquina (NUQ, 50), and Pampeana Colorada (PCA, 14); Bolivia – Criolla Boliviana (BOL, 40); Brazil – Canindé (CND, 40), Graúna (GRN, 40), Marota (MRT, 40), Moxotó (MOX, 40), Repartida (REP, 40), Azul (SAZ, 40) and undefined Creole (SRD, 36); Colombia – Criolla Colombiana (COL, 24); Cuba – Criolla Cubana (CUB, 40); Ecuador – Criolla del Ecuador (ECU, 11) and Galapagos Goat (GAG, 23); Paraguay – Criolla Paraguaya (PGY, 84); Peru – Criolla Peruana (PER, 61); United States of America – Myotonic Goat (MYO, 43), San Clemente Goat (SCL, 58), and Spanish Goat (SPA, 64); Venezuela – Criolla Venezolana (VEN, 46).

Bolivia – Criolla Boliviana (BOL, 40); Brazil – Canindé (CND, 40), Graúna (GRN, 40), Marota (MRT, 40), Moxotó (MOX, 40), Repartida (REP, 40), Azul (SAZ, 40) and undefined Creole (SRD, 36); Colombia – Criolla Colombiana (COL, 24); Cuba – Criolla Cubana (CUB, 40); Ecuador – Criolla del Ecuador (ECU, 11) and Galapagos Goat (GAG, 23); Paraguay – Criolla Paraguaya (PGY, 84); Peru – Criolla Peruana (PER, 61); United States of America – Myotonic Goat (MYO, 43), San Clemente Goat (SCL, 58) and Spanish Goat (SPA, 64); Venezuela – Criolla Venezolana (VEN, 46). Geographic locations of each population are shown in Fig. 1. Most of these populations are local breeds raised by small farmers, sometimes in family herds used for home consumption; but many are partially feral indistinct populations living free, as is the case of isolated GAG and SCL goat populations from Galapagos and San Clemente Islands respectively. The SRD is a poorly defined population of goats from Brazil resulting from admixture among several local breeds, possibly with influence from exotic germplasm. Samples were also collected from 112 animals of three commercial transboundary goat breeds, with samples collected in Brazil – Alpine (ALP, 35) and Anglo-Nubian (ANG, 41), and Spain – Saanen (SAAN, 36), to investigate admixture with local Creole goats. To minimize the degree of relationship among individuals, unrelated animals were selected whenever possible. For a broad representation of the genetic variability, samples were collected from several herds per breed. Biological samples were collected by qualified veterinarians during their routine practice in the framework of official health control programs. Therefore, no ethical approval was required for sampling of biological material. Genomic DNA was extracted from whole blood or

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

hair roots following the methodology described by Walsh *et al.* (1991).

STR genotyping

We analyzed a set of 21 short tandem repeat (STR) markers, selected following the recommendations of the ISAG-FAO Advisory Group on Farm Animal Genetic Diversity for studies in goats (Hoffmann *et al.* 2004; Hoffmann 2010), as follows: *BM1329*, *BM6506*, *BM6526*, *BM8125*, *CRSM60*, *CSRD247*, *ETH010*, *ETH225*, *HAUT27*, *ILSTS011*, *INRA063*, *MAF065*, *MAF209*, *McM527*, *MM12*, *OarFCB011*, *OarFCB048*, *OarFCB304*, *SPS115*, *SRCRSP08*, and *TGLA122*. STR markers were amplified in multiplex polymerase chain reactions (PCRs) using fluorescence-labeled primers as described by Bruno-de-Sousa *et al.* (2011). PCR fragments were separated by capillary electrophoresis on ABI 377XL instruments (Applied Biosystems), according to the manufacturer's recommendations. Allele sizes were determined using the internal size standard GeneScan-400HD ROX (Applied Biosystems), and genotypes were called using GENOTYPER® 2.5.1 software. Reference samples were included in each run for consistency of allele assignments.

Statistical analyses

Allele frequencies, effective number of alleles (n_e) and number of population-specific alleles (private alleles, PA) were determined with GENALEX v. 6.5 (Peakall & Smouse 2006) for each locus and breed. The frequency of null alleles (r) per locus within each breed was estimated with FREENA software (available at http://www1.montpellier.inra.fr/URLB/; Chapuis & Estoup 2007) following Dempster *et al.* (1977). Deviations from Hardy–Weinberg equilibrium (HWE) were assessed with GENEPOP v. 4 software (Raymond & Rousset 2003). Both global tests across populations and loci and tests per locus per population were carried out using the method of Guo & Thompson (1992), and the *P*-values were obtained using a Markov chain of 10 000 dememorization steps, 500 batches and 5000 iterations. Genotypic linkage disequilibrium (LD) was also calculated using this software and the same Markov chain settings. FSTAT v. 2.9.3 (Goudet 2001) was used to estimate the *F* statistics per locus according to Weir & Cockerham (1984), and *P*-values were obtained based on 1000 randomizations. Allelic richness (R_t) over all loci for each breed was also calculated using this software, assuming a minimum of three animals per breed.

The software GENETIX v. 4.05 (Belkhir *et al.* 1996–2004) was used to estimate observed (H_o) and unbiased within-breed expected (H_e) heterozygosities and mean number of alleles per breed (MNA) and to calculate the inbreeding coefficient (F_{is}) in each breed, with *P*-values obtained based on 1000 permutations. The 'plot.membership' function (Olivier François; http://membres-timc.imag.fr/Olivier.Francois/index.html) was used in R software v. 3.2.4 (1999–2016, R Foundation for Statistical Computing) to obtain contour graphics of geographical coordinates in relation to H_e and F_{is} values estimated for each Creole goat population.

A principal components analysis was carried out in PCAGEN v. 1.2 (available at http://www2.unil.ch/popgen/softwares/pcagen.htm; Goudet 1999) to represent breed relationships based on allele frequencies, assuming *P*-values for the axes obtained from 1000 randomizations of genotypes. Population pairwise D_A distances of Nei *et al.* (1983) were calculated in POPULATIONS (Langella 1999–2002) and used to obtain a network of breed relationships following the Neighbor-Net method in SPLITSTREE4 v. 4.13 (Huson & Bryant 2006).

Genetic structure and the degree of admixture of Creole native goats were investigated using the Bayesian clustering procedure of STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000). The most probable number of ancestral populations (K) given the observed genotypic data was estimated by performing six independent runs for each K ($1 < K < 30$) with burn-in lengths and Markov chain Monte Carlo (MCMC) iterations of 50 000 and 300 000 respectively. The alpha (degree of admixture) parameter was inferred from the data using the default settings and an admixture model with correlated allele frequencies (Falush *et al.* 2003). Assessment of the change in the log of the likelihood function and the method of Evanno *et al.* (2005) were used to identify the most probable K by determining the modal distribution of ΔK. The STRUCTURE analysis was repeated for some subsets of the data to assess within-breed substructure. Assignment tests were performed with STRUCTURE using prior information of source breeds and the settings described above for the MCMC. The proportion of each individual genotype in each cluster or breed (q) and the probability of ancestry in other breeds were estimated. The percentage of individuals assigned to source breeds was calculated for distinct threshold q values.

Results

Molecular markers

Among the 21 markers, there was no evidence for null alleles at important frequencies ($r > 0.2$) across various breeds, and a total of 236 alleles were detected at all loci across the various goat populations. Measures of genetic variability as well as *F*-statistics for each STR are shown in Table S1. The total number of alleles per locus ranged from 3 for *MAF209* to 21 for *OarFCB304*, and the mean number of alleles per locus ranged from 2.22 for *MAF209* to 9.52 for *BM6506*. The effective number of alleles, i.e. the number of alleles that if equally frequent would result in the observed homozygosity, was about four for most loci (*BM1329*, *CRSM60*, *CSRD247*, *OarFCB011*, *OarFCB048*,

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

OarFCB304, *SRCRSP08* and *TGLA122*), and ranged from one for *MAF209* to five for *BM6526*, *MAF065* and *MM12*. Allelic richness per locus, i.e. the number of different alleles independent of sample size, was also about four for most loci (*BM1329*, *BM6506*, *BM8125*, *CRSM60*, *CSRD247*, *ILSTS011*, *MAF065*, *McM527*, *OarFCB011*, *OarFCB048*, *OarFCB304*, *SRCRSP08* and *TGLA122*) and ranged from two for *ETH225*, *MAF209* and *SPS115* to five for *BM6526* and *MM12*. Private alleles (i.e., an allele at a given locus found exclusively in one population) were detected in 10 of the 21 loci studied; however, only for *ETH225*, *MAF065* and *MM12* were these private alleles detected in frequencies ≥ 0.05. The highest heterozygosity was found for *MM12* ($H_o = 0.760$ and $H_e = 0.768$) and *BM6526* ($H_o = 0.746$ and $H_e = 0.793$) and the lowest for *MAF209* ($H_o = 0.224$ and $H_e = 0.244$) and *ETH225* ($H_o = 0.273$ and $H_e = 0.319$). The F_{is} value per locus had an overall mean of 0.083 ± 0.058 and ranged from –0.012 for *BM8125* to 0.204 for *SPS115*. F_{it} and F_{st} per locus had overall means of 0.211 ± 0.058 and 0.134 ± 0.033 respectively. Out of the 21 STRs analyzed in each breed, *BM6506*, *MAF065*, *SPS115* and *SRCRSP08* showed significant ($P < 0.05$) HWE deviations across 10 or more breeds. The LD was significant ($P < 0.0001$) for 20 STR pairs; however, only the following three pairs appear to correspond to loci located in the same chromosome: *BM1329/SRCRSP08*, *BM8125/MAF209* and *BM8125/OarFCB048*.

Genetic diversity within populations

Estimates of within-breed genetic diversity are summarized in Table 1 and in the contour graphics shown in Fig. 1. Overall, the genetic diversity of the 24 Creole goat populations analyzed here (MNA = 5.82 ± 1.14, $n_e = 3.30 \pm 0.67$, $R_t = 3.01 \pm 0.36$, $H_o = 0.585 \pm 0.074$ and $H_e = 0.638 \pm 0.079$) was moderate and slightly lower than that detected with a similar panel of STRs in their European counterparts (see Cañon *et al.* 2006; Bruno-de-Sousa *et al.* 2011). Isolated Creole goat populations from Galapagos and San Clemente Islands had the lowest diversity with respectively 3.05 ± 1.24 and 4.14 ± 1.39 for MNA, 1.98 ± 0.69 and 2.10 ± 0.89 for n_e, 2.12 ± 0.64 and 2.24 ± 0.65 for R_t, 0.411 ± 0.258 and 0.389 ± 0.183 for H_o, and 0.433 ± 0.243 and 0.452 ± 0.208 for H_e (green areas in Fig. 1a). The highest genetic diversity was found in PGY from Paraguay and SPA from North America with respectively 8.71 ± 3.32 and 7.81 ± 2.98 for MNA, 5.19 ± 2.02 and 4.24 ± 1.71 for n_e, 3.73 ± 0.68 and 3.43 ± 0.77 for R_t, 0.677 ± 0.132 and 0.678 ± 0.168 for H_o, and 0.777 ± 0.101 and 0.717 ± 0.153 for H_e respectively (pink areas in Fig. 1a). Despite their low sample size, the Argentinean longhair goat populations TAN and CHL had high genetic diversity: $H_o = 0.668 \pm 0.173$ and $H_o = 0.669 \pm 0.160$, and $H_e = 0.711 \pm 0.124$ and $H_e = 0.720 \pm 0.130$ respectively. Also, the crossbred SRD goat population from Brazil showed high genetic diversity (MNA = 6.33 ± 2.83, $n_e = 3.87 \pm 1.47$, $R_t = 3.28 \pm 0.79$, $H_o = 0.640 \pm 0.182$ and $H_e = 0.697 \pm 0.178$). Among the three commercial transboundary breeds analyzed, ALP had the highest diversity (MNA = 6.76 ± 2.72, $n_e = 4.08 \pm 1.65$, $R_t = 3.37 \pm 0.83$, $H_o = 0.674 \pm 0.205$ and $H_e = 0.701 \pm 0.198$). Overall, the genetic diversity estimates for this breed group were similar to those observed for Creole goats (MNA = 6.43 ± 0.50, $n_e = 3.55 \pm 0.53$, $R_t = 3.15 \pm 0.22$, $H_o = 0.634 \pm 0.036$ and $H_e = 0.664 \pm 0.033$). Private alleles were detected in 14 breeds (Argentinean CHC and NUQ; Brazilian GRN, MRT, MOX and SAZ; Ecuadorian ECU and GAG; Paraguayan PGY; Peruvian PER; North American MYO; Venezolana VEN; and transboundary ALP and SAAN); however, only three out of 17 PAs were detected in frequencies ≥ 0.05 (one each in MYO, ECU and CHC). The number of loci that had a significant heterozygote deficit ($P < 0.05$) ranged from 0 (PCA) to 14 (PGY). The HWE deviations detected are most probably due to inbreeding, as the within-breed F_{is} estimates were significantly ($P < 0.05$) higher than zero in all populations, except in Argentinean TAN, NUQ and PCA; Brazilian GRN; and Galapagos Goat (depicted in green in Fig. 1b). For 13 Creole goat populations, the $F_{is} > 0$ was highly significant ($P < 0.001$). Heterozygote excess was not significant for any of the locus/breed combinations tested.

Genetic relationship among populations

Pairwise population F_{st} values, as shown in Table S2, were mostly significant for $P < 0.05$. The lowest F_{st} values were found among Argentinean goat populations (overall mean of 0.041 ± 0.025; Table S3), particularly between CHL and CHC (–0.008), PCA (0.017) and TAN (0.020), CHC and TAN (0.021) and PCA and TAN (0.027), which were not significantly differentiated ($P < 0.05$). In spite of being significant, F_{st} values among Brazilian breeds were also low (0.079 ± 0.022), namely of MOX with CND (0.038) and REP (0.046), SAZ with GRN (0.040) and CND with REP (0.050). The highest F_{st} value was observed between Creole goat populations from Galapagos and San Clemente Islands (0.391). The overall F_{st} value among the 24 Creole and three commercial transboundary breeds was 0.134 ± 0.071 (Table S3). The lowest F_{st} values between Creole and commercial transboundary goats were found for SPA/SAAN (0.054) and SPA/ALP (0.055), whereas the highest was observed for GAG/ANG (0.299).

Breed relationships based on allele frequencies are depicted in the principal components analysis graph in Fig. 2. The island breeds GAG and SCL were the most distant, together with BOL and the commercial Anglo-Nubian goats. Based on axes 1 and 2, Brazilian breeds clustered together (shaded area in gray) and split from all other goat populations, which had a more central position. The Argentinean breeds also grouped together such that,

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Table 1 Mean number of alleles (MNA), effective number of alleles (n_e), allelic richness (R_t), number of private alleles (PA), observed (H_o) and expected (H_e) heterozygosities, inbreeding coefficient (F_{is}), number of loci showing heterozygote deficit (HWEd) estimated with 21 STRs in 24 Creole and three commercial transboundary goat breeds.

Country	Breed	Abr.	*n*	MNA ± SD	n_e ± SD	R_t ± SD	PA	H_o ± SD	H_e ± SD	F_{is}	HWE
Creole											
Argentina	Angora	TAN	14	5.62 ± 2.50	3.74 ± 1.83	3.28 ± 0.84		0.668 ± 0.173	0.711 ± 0.124	0.062	2
Argentina	Chilluda longhair	CHL	12	5.95 ± 2.31	3.76 ± 1.69	3.40 ± 0.77		0.669 ± 0.160	0.720 ± 0.130	0.073*	2
Argentina	Chilluda shorthair	CHC	11	5.43 ± 2.58	3.73 ± 1.87	3.29 ± 0.92	**1**	0.650 ± 0.196	0.696 ± 0.179	0.069*	2
Argentina	Cr. del Nordeste	NEA	39	6.43 ± 2.60	3.77 ± 1.47	3.24 ± 0.76		0.596 ± 0.181	0.690 ± 0.166	0.138***	9
Argentina	Neuquina	NUQ	50	6.76 ± 2.14	3.32 ± 1.21	3.09 ± 0.72	1	0.643 ± 0.226	0.648 ± 0.188	0.008	5
Argentina	Pampeana Colorada	PCA	14	4.86 ± 1.96	3.28 ± 1.44	3.09 ± 0.79		0.643 ± 0.205	0.664 ± 0.153	0.032	0
Bolivia	Cr. Boliviana	BOL	40	5.24 ± 2.28	3.21 ± 1.18	2.97 ± 0.70		0.570 ± 0.159	0.648 ± 0.154	0.121***	9
Brazil	Canindé	CND	40	5.52 ± 2.09	3.11 ± 1.10	2.94 ± 0.72		0.589 ± 0.187	0.634 ± 0.175	0.071***	3
Brazil	Graúna	GRN	40	6.24 ± 2.39	3.07 ± 1.24	2.92 ± 0.72	1	0.603 ± 0.199	0.620 ± 0.187	0.027	4
Brazil	Marota	MRT	40	5.24 ± 2.19	2.75 ± 1.08	2.73 ± 0.69	1	0.555 ± 0.183	0.590 ± 0.170	0.061**	6
Brazil	Moxotó	MOX	40	5.52 ± 2.14	2.93 ± 1.09	2.79 ± 0.75	1	0.524 ± 0.213	0.604 ± 0.201	0.133***	6
Brazil	Repartida	REP	40	6.14 ± 2.33	3.10 ± 1.18	2.97 ± 0.70		0.589 ± 0.193	0.632 ± 0.166	0.068***	3
Brazil	Azul	SAZ	40	5.52 ± 2.29	2.61 ± 1.17	2.61 ± 0.76	1	0.517 ± 0.232	0.547 ± 0.214	0.056**	2
Brazil	Undefined Creole	SRD	36	6.33 ± 2.83	3.87 ± 1.47	3.28 ± 0.79		0.640 ± 0.182	0.697 ± 0.178	0.082***	6
Colombia	Cr. Colombiana	COL	24	5.90 ± 1.89	3.63 ± 1.33	3.24 ± 0.78		0.592 ± 0.217	0.680 ± 0.196	0.132***	6
Cuba	Cr. Cubana	CUB	40	6.00 ± 2.14	3.03 ± 1.14	2.92 ± 0.65		0.574 ± 0.173	0.629 ± 0.153	0.088***	5
Ecuador	Cr. Del Ecuador	ECU	11	4.57 ± 1.83	2.79 ± 1.18	2.84 ± 0.88	**1**	0.542 ± 0.298	0.587 ± 0.237	0.081*	3
Ecuador	Galapagos goat	GAG	23	3.05 ± 1.24	1.98 ± 0.69	2.12 ± 0.64	1	0.411 ± 0.258	0.433 ± 0.234	0.053	3
Paraguay	Cr. Paraguaya	PGY	84	8.71 ± 3.32	5.19 ± 2.02	3.73 ± 0.68	3	0.677 ± 0.132	0.777 ± 0.101	0.130***	14
Peru	Cr. Peruana	PER	61	6.86 ± 2.69	3.45 ± 1.55	3.05 ± 0.85	1	0.561 ± 0.223	0.642 ± 0.189	0.126***	11
USA	Myotonic goat	MYO	43	6.33 ± 2.18	3.31 ± 1.24	3.06 ± 0.66	**1**	0.580 ± 0.175	0.661 ± 0.139	0.124***	7
USA	San Clemente goat	SCL	58	4.14 ± 1.39	2.10 ± 0.89	2.24 ± 0.65		0.389 ± 0.183	0.452 ± 0.208	0.139***	4
USA	Spanish goat	SPA	64	7.81 ± 2.98	4.24 ± 1.71	3.43 ± 0.77		0.678 ± 0.168	0.717 ± 0.153	0.056***	6
Venezuela	Cr. Venezolana	VEN	46	5.38 ± 2.11	3.15 ± 1.21	2.93 ± 0.80	1	0.591 ± 0.217	0.623 ± 0.206	0.053**	3
Means			910	5.82 ± 1.14	3.30 ± 0.67	3.01 ± 0.36		0.585 ± 0.074	0.638 ± 0.079	0.083	121
Commercial transboundary											
Brazil	Alpine	ALP	35	6.76 ± 2.72	4.08 ± 1.65	3.37 ± 0.83	1	0.674 ± 0.205	0.701 ± 0.198	0.040*	3
Brazil	Anglo-nubian	ANG	41	5.86 ± 2.13	3.01 ± 0.95	2.94 ± 0.58		0.604 ± 0.140	0.640 ± 0.128	0.057*	2
Spain	Saanen	SAAN	36	6.67 ± 2.37	3.57 ± 1.57	3.15 ± 0.86	2	0.624 ± 0.248	0.650 ± 0.225	0.041*	3
Means			112	6.43 ± 0.50	3.55 ± 0.53	3.15 ± 0.22		0.634 ± 0.036	0.664 ± 0.033	0.046	8
Overall			1022	5.88 ± 1.08	3.32 ± 0.64	3.02 ± 0.34	17	0.591 ± 0.071	0.640 ± 0.074	0.083	129

Breeds are grouped according to geographic area.

Abr., breed abbreviation; *n*, sample size; *SD*, standard deviation; PA, private alleles, frequency >5% in bold; HWE, Hardy-Weinberg equilibrium for $P < 0.05$.

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

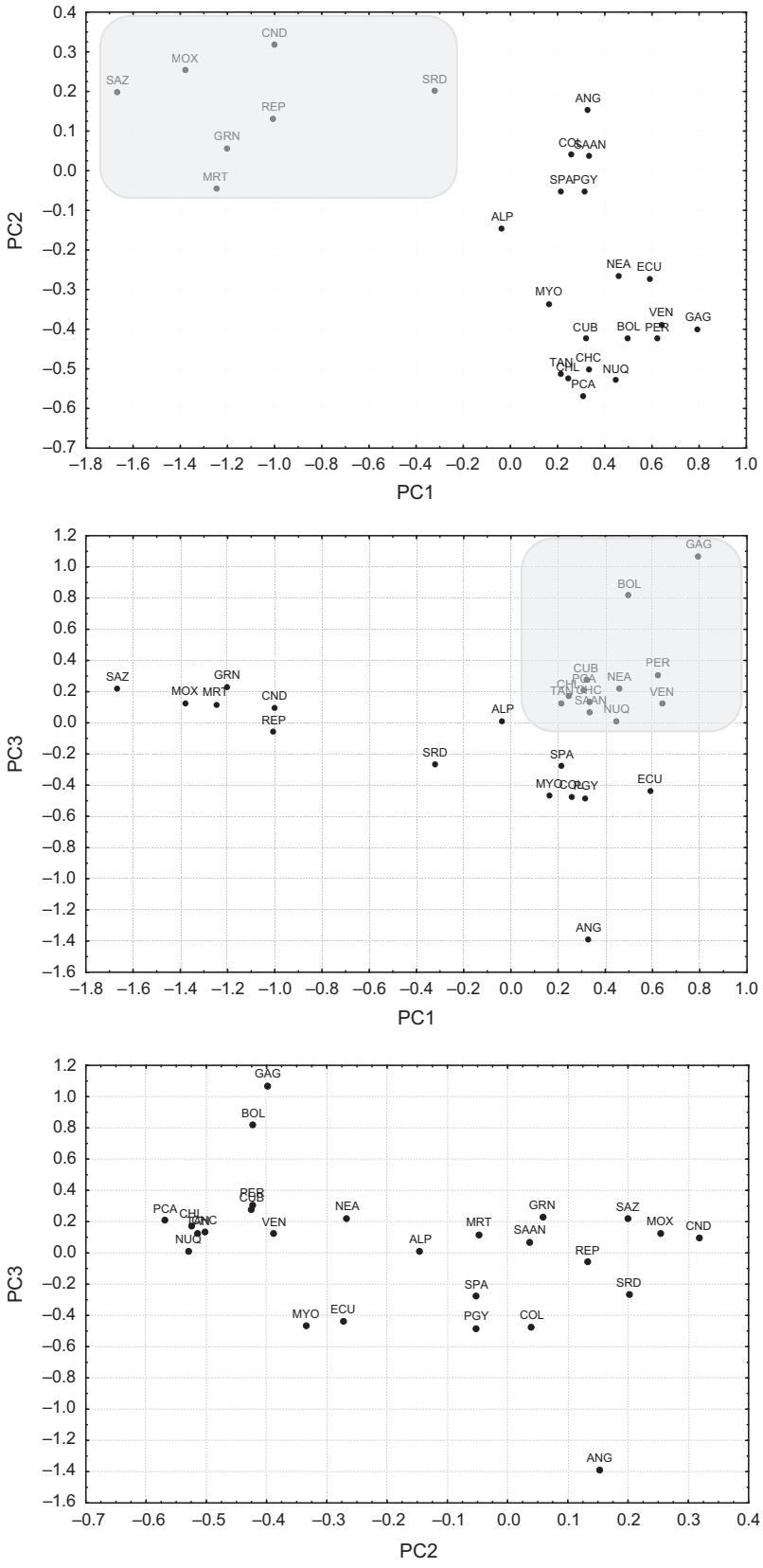


Figure 2 Analysis of principal components graph depicting breed relationships among 24 Creole goat populations and three commercial transboundary goat breeds based on allele frequencies. Breed codes are defined in Table 1. Shaded areas in gray depict groups of more closely related breeds. Axis 1: 25.15% inertia, $P = 0.001$; Axis 2: 15.86% inertia, $P = 0.001$; Axis 3: 8.91% inertia, $P = 0.01$.

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

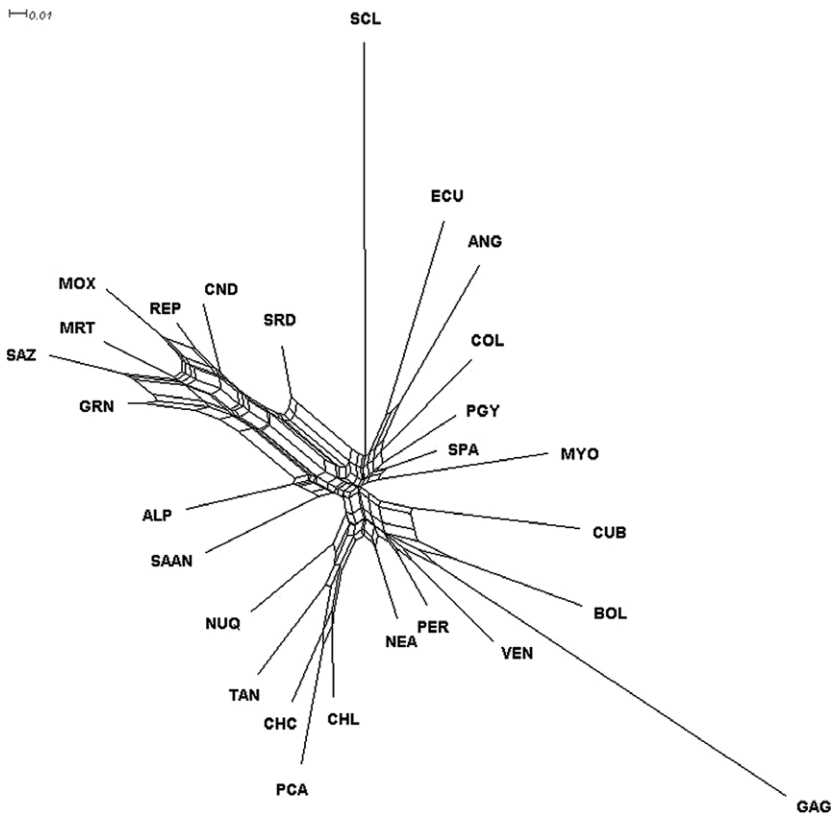


Figure 3 Neighbor-Net graph of D_A genetic distances depicting breed relationships among 24 Creole goat populations and three commercial transboundary goat breeds.

based on axes 1 and 3, they had a close genetic relationship with the Creole CUB, PER and VEN and the commercial Saanen goats (shaded area in gray) and split from MYO, SPA, COL, PGY and ECU.

The Neighbor-Net of D_A genetic distances is shown in Fig. 3, and pairwise population distance values are shown in Table S2. The Brazilian goat breeds formed a tight net (supported by a high bootstrap value of 94%; result not shown) and were somewhat genetically close to the transboundary ALP and SAAN breeds. The Argentinean Creole goats formed another breed group (bootstrap value of 75%; result not shown) somewhat related to the PER, VEN, GAG, BOL and CUB. Criolla del Ecuador clustered with the commercial ANG breed and close to COL, PGY and SPA and, to a lesser extent, with the MYO breed. The San Clemente population split from the center as an independent and more distant branch.

Genetic structure and admixture analysis

For the STRUCTURE analysis done without including prior information on sample origin and following Evanno *et al.* (2005), the highest ΔK values were found for $K = 4$ ($\Delta K = 91.1$) and $K = 21$ ($\Delta K = 74.3$) (Fig. S1). The structure detected at $K = 4$ corresponds to the lowest variance of the likelihood of the data estimated over six independent runs, whereas the highest likelihood of the data was obtained for $K = 21$. The graphical representation of individual genotype membership coefficients (q) in each cluster is shown in Fig. 4. Based on individual q values and at $K = 4$, the Brazilian breeds grouped together (except for SRD, which showed influence of the commercial breeds, as expected), as did the Argentinean breeds, which were grouped with other more closely related Creole breeds (CUB, VEN, GAG, PER, BOL; Fig. 4a), in agreement with the results of the Neighbor-Net (Fig. 3). Influence of the commercial transboundary breeds in Creole goats could be inferred for SPA, MYO, COL, ECU and PGY, whereas the isolated SCL goat population formed an independent cluster. The breed groups at $K = 21$ (Fig. 4b, Table 2) indicate that most breeds formed independent clusters, whereas PGY split into two groups and the following breeds were grouped together: Brazilian SAZ and GRN; Brazilian MOX, CND and REP; Argentinean CHL, CHC, PCA and TAN; Ecuadorian ECU; and the North American SPA.

Further analyses were carried out with subsets of the data to determine if individuals of the Brazilian and Argentinean breed groups could be distinguished. The highest values of ΔK were obtained at $K = 3$ ($\Delta K = 287.7$) and $K = 5$ ($\Delta K = 72.2$) for the analysis of the Brazilian breeds SAZ, MOX, MRT, CND, REP and GRN (SRD was not included because it is a poorly defined and highly crossbred population). ALP was included as a reference breed for this analysis because it is a well-differentiated population and thus useful for determining the most probable K value of this data subset. Based on individual q values and at $K = 3$,

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

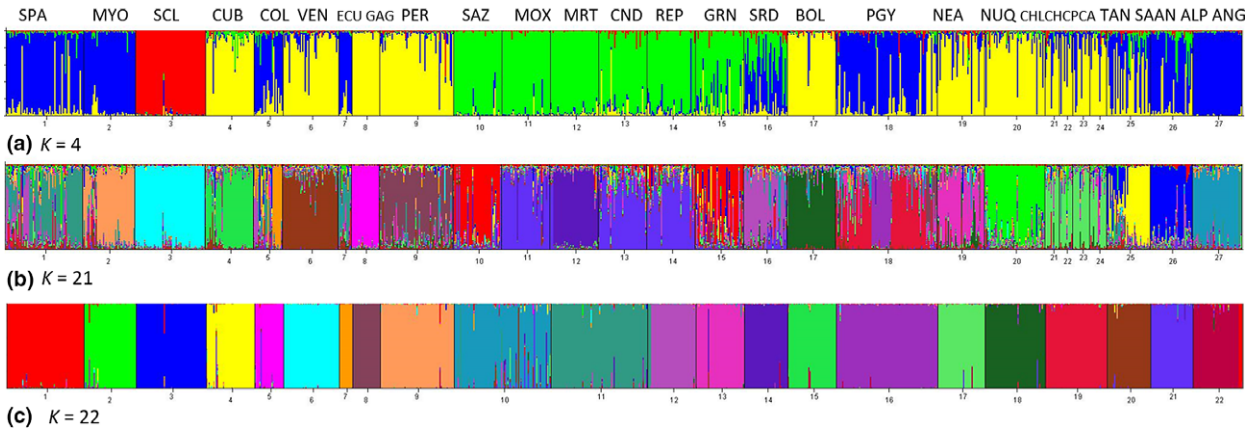


Figure 4 Graphical representation of individual genotype membership coefficients (q) in each cluster obtained with STRUCTURE software without prior information on sample origin for (a) $K = 4$ and (b) $K = 21$ and (c) considering prior information on source breeds for $K = 22$.

the SAZ, MRT and GRN breeds formed one group; MOX, CND and REP another; and ALP an independent cluster (Fig. S2). At $K = 5$, MRT and CND formed independent clusters, whereas MOX and REP were in the same group as were SAZ and GRN. For the subset of Argentinean goats, the modal value of $\Delta K = 129.8$ was obtained at $K = 4$ with breeds NEA, NUQ and the reference ALP each in a separate cluster and CHL, CHC, PCA and TAN in the same group (Fig. S3).

The results of the Bayesian analysis carried out with STRUCTURE are summarized in Table 2, where the mean membership coefficients and percentages of animals assigned to each cluster are shown for scenarios without and with prior information on sample origin ($K = 21$ or $K = 22$ respectively). When prior information on sample origin was ignored, and assuming $K = 21$, the average membership proportions in each cluster (Q) ranged from 0.447 ± 0.208 in ECU to $0.937 \pm 0.0.035$ in GAG with an overall average of 0.735 ± 0.207. Approximately 63% of the individuals were classified within their source cluster assuming a threshold of $q \geq 0.700$, whereas for more stringent threshold q values, only ~25% ($q \geq 0.900$), ~7% ($q \geq 0.950$) and 0% ($q \geq 0.999$) of the animals were correctly assigned (result not shown). There were several heterogeneous breeds (SPA, COL, ECU, PER, GRN, REP, CND, SRD, PGY, NEA, CHC and TAN) with less than 60% of the individuals assigned to their source cluster (for $q \geq 0.700$) and/or over 30% of crossbred animals (i.e. with individual genotype membership q values < 0.700 but maximum in the source cluster). Also, the relatively high percentage of crossbred animals observed in the ALP breed (~29%) appears to reflect the influence of this breed in some Brazilian Creole goats (10 ALP animals showed $q \geq 10\%$ within the Brazilian breed group). Likewise, six ANG animals had $q \geq 10\%$ in the SPA/ECU cluster (and the ECU breed had ~13% of its mean genotype membership coefficient in the ANG cluster), which indicates some level of admixture between these breeds. For a threshold of $q \geq 0.700$, approximately 2% of the animals were misclassified to a cluster other than their source, with SAZ from Brazil showing 10% of the animals assigned to a different breed. The highest percentage of animals assigned to the source cluster (100% with $q \geq 0.700$) was observed in the GAG population, followed by six other breeds (SCL, VEN, MOX, MRT, BOL and PCA) with >78% of individuals assigned for the same stringency threshold.

When prior information on source breeds was considered in the analysis (i.e. assuming the 22 breed clusters obtained in STRUCTURE without priors), the overall Q was 0.919 ± 0.212 and ranged from 0.605 ± 0.464 in GRN to 0.998 ± 0.004 in GAG (Table 2). Individual q values for $K = 22$ are depicted in Fig. 4c. The overall percentage of individuals assigned to the source cluster or breed was about 92% for $q \geq 0.700$ and 81% for $q \geq 0.950$, but this value fell considerably with the highest stringency (~17% with $q \geq 0.999$). The percentage of animals assigned to the source cluster with $q \geq 0.950$ was greater than 70% in all breeds except GRN (50%), REP (65%) and CHL (67%). Among the group of Creole goats analyzed, GAG (100%), SCL (97%), VEN (94%), PCA (93%), ECU (91%) and MOX (90%) had the highest percentage of individuals assigned with $q \geq 0.950$. In fact, for the GAG and SCL populations, there were about 78% of animals with $q \geq 0.999$. Several of the animals that showed evidence of admixture were assigned to their source cluster ($q \geq 0.950$) when prior information was considered. However, and particularly in the case of SAZ, REP, MRT and GRN, admixture was detected within the Brazilian breed group (including SRD goats) even when the source clusters were predefined. In the case of GRN, there were 14 individuals that showed ancestry across several source populations, including other Creole and commercial transboundary breeds (SAAN and ALP).

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Table 2 Summary results of the Bayesian cluster analysis done with STRUCTURE both with and without prior information on sample origin. Average genotype membership proportions (Q), the percentage of animals assigned to each cluster considering various threshold *q* values, crossbred (Cross.) and misclassified (Miscl.) individuals at the lowest stringency (*q* values of 0.700) are shown for each breed and overall.

	Without prior information on sample origin							With prior information on sample origin					
			% assigned to each cluster							% assigned to each cluster			
Breeds	Cluster	Q ± SD	$q \geq 0.700$	$q \geq 0.900$	$q \geq 0.950$	% Cross.[1] $q < 0.700$	% Miscl.[2] $q \geq 0.700$	Cluster	Q ± SD	$q \geq 0.700$	$q \geq 0.900$	$q \geq 0.950$	$q \geq 0.999$
Creole													
SPA	11	0.595 ± 0.228	42.2	3.1	0.0	25.0	4.7	1	0.960 ± 0.062	98.4	90.6	85.9	0.0
MYO	9	0.714 ± 0.258	74.4	25.6	2.3	18.6		2	0.898 ± 0.245	93.0	83.7	76.7	14.0
SCL	6	0.920 ± 0.076	98.3	77.6	46.6	1.7		3	0.980 ± 0.131	98.3	98.3	96.6	77.6
CUB	15	0.778 ± 0.177	75.0	27.5	7.5	15.0		4	0.888 ± 0.272	90.0	82.5	82.5	25.0
COL	7	0.602 ± 0.257	33.3	29.2	12.5	12.5	4.2	5	0.891 ± 0.215	83.3	79.2	75.0	12.5
VEN	20	0.822 ± 0.120	84.8	32.6	0.0	15.2		6	0.983 ± 0.036	100.0	95.7	93.5	8.7
ECU	11	0.447 ± 0.208	9.1	0.0	0.0	54.5		7	0.977 ± 0.022	100.0	100.0	90.9	9.1
GAG	5	0.937 ± 0.035	100.0	87.0	47.8	0.0		8	0.998 ± 0.004	100.0	100.0	100.0	78.3
PER	8	0.690 ± 0.191	55.7	9.8	0.0	37.7		9	0.947 ± 0.140	95.1	90.2	82.0	1.6
SAZ	1	0.794 ± 0.166	67.5	35.0	12.5	12.5	10.0	10	0.825 ± 0.322	77.5	72.5	72.5	32.5
GRN	1	0.662 ± 0.261	47.5	30.0	10.0	30.0	5.0	10	0.605 ± 0.464	60.0	55.0	50.0	30.0
MOX	21	0.814 ± 0.171	82.5	40.0	5.0	10.0	2.5	11	0.935 ± 0.203	92.5	90.0	90.0	15.0
REP	21	0.693 ± 0.185	57.5	7.5	0.0	32.5		11	0.853 ± 0.271	82.5	77.5	65.0	5.0
MRT	14	0.852 ± 0.112	87.5	37.5	2.5	5.0		12	0.899 ± 0.280	90.0	90.0	85.0	17.5
CND	21	0.677 ± 0.180	52.5	0.0	0.0	40.0		13	0.929 ± 0.190	92.5	87.5	80.0	7.5
SRD	12	0.650 ± 0.205	44.4	8.3	0.0	36.1		14	0.956 ± 0.117	97.2	94.4	88.9	0.0
BOL	18	0.837 ± 0.150	87.5	42.5	2.5	12.5		15	0.928 ± 0.227	92.5	92.5	87.5	22.5
PGY	19; 16	0.669 ± 0.202	36.9	2.4	0.0	27.4	1.2	16	0.959 ± 0.071	97.6	91.7	79.8	0.0
NEA	13	0.644 ± 0.198	43.6	7.7	0.0	41.0		17	0.940 ± 0.178	94.9	92.3	89.7	2.6
NUQ	2	0.747 ± 0.192	66.0	26.0	4.0	32.0		18	0.934 ± 0.177	94.0	84.0	78.0	14.0
CHL	17	0.710 ± 0.188	66.7	8.3	0.0	16.7		19	0.933 ± 0.103	91.7	83.3	66.7	8.3
CHC	17	0.664 ± 0.149	36.4	0.0	0.0	45.5		19	0.959 ± 0.054	100.0	90.9	72.7	0.0
PCA	17	0.809 ± 0.202	78.6	35.7	7.1	14.3		19	0.988 ± 0.017	100.0	100.0	92.9	14.3
TAN	17	0.671 ± 0.169	42.9	0.0	0.0	42.9	7.1	19	0.946 ± 0.088	92.9	85.7	71.4	0.0
Transboundary													
SAAN	4	0.678 ± 0.237	50.0	25.0	5.6	8.3	8.3	20	0.935 ± 0.137	94.4	83.3	75.0	11.1
ALP	3	0.756 ± 0.166	65.7	11.4	0.0	28.6		21	0.948 ± 0.130	94.3	88.6	82.9	2.9
ANG	10	0.813 ± 0.146	73.2	39.0	17.1	19.5	7.3	22	0.859 ± 0.296	85.4	73.2	73.2	36.6
Overall	21	0.735 ± 0.207	62.7	24.5	6.8	22.6	1.9	22	0.919 ± 0.212	92.1	86.9	81.3	16.7

[1]Crossbred: animals with individual genotype membership *q* values < 0.700 but maximum in the source cluster.
[2]Misclassified: animals with $q \geq 0.700$ in a cluster other than the source.

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Discussion

This study covered a wide representation of goat genetic diversity in the Americas, and it represents the first comprehensive analysis of the five-century history of Creole goats. The microsatellite-based analysis adopted here accounts for the direct influence of genetic drift and migration in the evolution of Creole goat populations. Indirectly, it could also reflect the effects of natural and human selection in shaping the genetic composition of these local and well-adapted breeds. Our results depict the genetic diversity and population structure of 24 Creole goat populations from 10 countries and also the possible influence of three commercial transboundary breeds of worldwide distribution, which is relevant for the management and conservation of these important farm animal genetic resources. The genetic markers chosen follow FAO recommendations for genetic diversity studies, with a total number of alleles greater than four (except for *ETH225*, *MAF209* and *SPS115*) and no null alleles detected at significant frequencies. The 21 STR markers used detected high levels of genetic diversity, with a total of 236 alleles observed across the 27 goat populations analyzed, and proved efficient for investigating the genetic relationships and population structure of Creole goats.

We observed an important level of within-breed genetic diversity in the Creole goats tested (means for $H_o = 0.585$ and $H_e = 0.638$). Although it is difficult to make comparisons with other studies that used different STR markers, the overall genetic diversity of Creole goats was analogous to that of their Central and North European counterparts ($H_o = 0.594$ and $H_e = 0.654$) but lower than that estimated for East and Central Mediterranean goats ($H_o = 0.663$ and $H_e = 0.737$), as reported by Cañon *et al.* (2006). When compared to the presumed ancestral goat breeds of the Iberian Peninsula ($H_o = 0.610$ and $H_e = 0.650$), Creole populations show slightly lower levels of genetic diversity than do their Western European counterparts (Martinez *et al.* 2015). The somewhat low mean values of parameters assessing the overall diversity of Creole goats, particularly for allelic richness ($R_t = 3.01$) can be explained, in part, as the repercussion of the extremely low diversity found in breeds from San Clemente and Galapagos islands ($R_t \approx 2.1$ and $H_o \approx 0.400$), which is undoubtedly a consequence of genetic drift and inbreeding. But it can be expected that for other Creole breeds a strong founder effect and subsequent bottleneck events may have occurred, as few animals originating from the Iberian Peninsula reached the Antilles in the 15th century and had an exponential growth afterwards. Subsequent founder effects can probably be associated with the selection of some of these animals to colonize the continent in three directions (Primo 1992): to Veracruz (Mexico) and from there towards the north of the continent; to Panama and then towards Central America and the northern region of the Vice Kingdom of Peru; and from southern Peru towards Río de La Plata (although an alternative route of goat introduction in Patagonia by sea through Valdivia in Chile is considered; see Bandieri *et al.* 1993). Also, the Portuguese used the so-called *capitanias* to introduce goats in Brazil along the coast. Despite the heterogeneous origins of Creole goats, over time a few founder animals dispersed and were used to create new breeds well adapted to the distinct and often extreme environments of the Americas. Since the 19th century, the replacement of some Creole goat populations by transboundary commercial breeds, following model breeding programs developed in Europe, resulted in the current threatened status of many populations. This scenario can further explain the lower diversity detected here when compared to goats from other regions, such as the Middle East (Cañon *et al.* 2006; Agha *et al.* 2008; Ağaoğlu & Ertuğrul 2012; Mahmoudi *et al.* 2014), Asia (Li *et al.* 2002; Fatima *et al.* 2008; Dixit *et al.* 2012) and West Africa (Traoré *et al.* 2009; Missohou *et al.* 2011). The few studies on Latin American goats reported similar overall values for the diversity parameters estimated in our study (Menezes *et al.* 2006; Oliveira *et al.* 2010; Calvo *et al.* 2012; Chacón *et al.* 2012; Ribeiro *et al.* 2012; Aranguren-Méndez *et al.* 2013). We observed that, for several Creole goat populations, there was a large number of loci showing deviations from HWE (i.e. heterozygote deficit) along with significant F_{is} estimates (in some cases >12%) which can be due to inbreeding and/or within-breed substructure, i.e. a *Wahlund* effect (Hartl & Clark 1997) associated with subdivided populations (for example, in some Creole breeds it is common to find large isolated herds).

Our results provide a broad perspective on the extant genetic diversity and population structure of Creole goats and to some extent can be interpreted as reflecting the mode of goat colonization in the Americas and the subsequent dilution resulting from crossbreeding with transboundary commercial breeds observed in a few Creole populations. We found some patterns of clear geographical structure, with strong consistency between genetic distances among breeds and the Bayesian STRUCTURE analyses. For example, the Neighbor-Net graph showed a clear separation of the Brazilian and Argentinean breed groups relative to the remaining populations, which was supported by the clusters defined with STRUCTURE. The average F_{st} estimates also support a high level of breed differentiation among Creole goat populations, even when the highly divergent San Clemente and Galapagos island goats are excluded from the analyses, with an estimated $F_{st} = 11\%$, which is larger than the overall F_{st} value of 7% reported by Cañon *et al.* (2006) for Northern European, Mediterranean and Middle Eastern goat breeds.

The observed level of genetic differentiation among goat populations from the various Latin American regions could possibly be associated with the two distinct colonization routes of the Spanish (Rodero *et al.* 1992) and the

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Portuguese conquerors (Primo 1992), and further studies including Iberian breeds are essential to clarify this aspect. Under this perspective, the yellow cluster defined with STRUCTURE for $K = 4$ (supported by the highest ΔK value) could represent the original Spanish colonization route, grouping Cuban goats (as putative representatives of those initially introduced in the Caribbean Islands) with Creole goats from Venezuela, Galapagos, Peru, Bolivia and Argentina (following the migratory route of Rio de La Plata towards Patagonia). In Creole goats from Colombia, Ecuador and Paraguay, the more recent influence of commercial transboundary breeds was detected, namely that of the Anglo-Nubian, which is a dairy goat breed better adapted to the tropics, and possibly diluting former signatures of Iberian ancestry. Such influence was also observed in the SPA and MYO breeds of the USA within the northern route of goat dispersion in the Americas. Creole goats from Argentina can be separated in two breed groups by region: the northeastern (NEA) breed, and the central (PCA) and southern breeds (CHL, CHC and NUQ). The NEA is entangled with the Creole goats from other somewhat neighbouring South American countries, such as Peru and Bolivia, but also with the breeds from Venezuela and Cuba. Within the central-southern breed group, the PCA (from the western region of La Pampa province) clustered together with the Creole Chilluda goats (CHL and CHC) and the Angora-type goats, whereas the NUQ from the region of Patagonia formed an independent cluster, possibly because of founder effects and genetic isolation. The somewhat low mean F_{st} observed for the Argentinean breed group (~4%) suggests that CHL and CHC indeed represent the same breed.

The differentiation of the Brazilian breeds, depicted in green in Fig. 4, may reflect their common origin from goats introduced in the 16th century by the Portuguese in the northeastern region. However, Ribeiro *et al.* (2012) detected some level of genetic proximity only between the GRN and CND breeds and their Portuguese counterparts. Yet, the possible influence of goat populations from other regions, namely from West Africa or Asia, cannot be discarded, as these could also have reached the Americas via trade routes. Indeed, the STRUCTURE analysis showed clear evidence that GRN is a heterogeneous population, for which influences from breeds not included in this analysis could explain its low percentage of individuals assigned. The substructure analysis revealed three major breed groups in Brazil, as follows: the MOX breed, which is white or cream with black legs and a black stripe on the back (a similar pattern to their Portuguese counterpart Serpentina), clustered with REP, whose name refers to the split colour between the black forequarters and brown, pale hind (or vice versa); the white-coloured MRT clustered with the CND breed, which is black with a white belly; and the black-coloured GRN, of putative Iberian ancestry (Ribeiro *et al.* 2012), clustered with the blue-coloured SAZ goats, allegedly of West African origin but also with phenotypic patterns similar to that shown by the Portuguese Serrana breed.

Despite the typical genetic heterogeneity observed in goats, the assignment results obtained in our study were powerful, with seven breeds (GAG, SCL, VEN, MOX, MRT, BOL and PCA) showing more than 78% of individuals assigned to their source cluster when the analyses ignored prior information on sample origin ($q \geq 0.700$). It is not surprising that these corresponded to naturally isolated breeds or to those with a more organized breeder's structure. Nevertheless, and except for three populations (GRN, REP and CHL), all Creole breeds had over 70% of individuals assigned when a threshold q value of 0.950 was adopted. These results indicate that most Creole goat breeds have their own identity, which supports these breeds as targets of management and conservation programs. These results also provide a first basis for the establishment of breed registries (i.e. herdbooks), which are still lacking in most cases.

Our study consisted of a fine-scale analysis of the genetic structure of Creole goats, but a more comprehensive scrutiny of worldwide goats is needed to unveil the origins of Creole goats of the Americas and the multiple influences that they have suffered, namely by including Iberian and West African populations along with other widely dispersed commercial transboundary breeds.

Groeneveld *et al.* (2010) concluded in their general review of farm animal diversity studies that further molecular analyses of autosomal and Y-chromosomal genetic variation are necessary to reveal the evolutionary trajectories of goats. Recently, high-density SNP arrays have been developed for several livestock species, including goats (Tosser-Klopp *et al.* 2014), and the genotyping cost has dropped significantly, making it possible to carry out genetic analysis at the population level with a large number of loci. These whole-genome markers as well as next-generation genome sequencing will allow for more refined biodiversity analyses in farm animals (Lenstra *et al.* 2012; Ajmone-Marsan *et al.* 2014) to investigate adaptation traits (Bruford *et al.* 2015) and to disclose evolutionary signatures that will contribute to a better understanding of their domestication histories. In goats, SNP arrays have been used for parentage testing (Talenti *et al.* 2016), to investigate genetic diversity and population structure (Nicoloso *et al.* 2015; Visser *et al.* 2016) and to detect selection signatures (Burren *et al.* 2016; Kim *et al.* 2016). However, whole-genome SNP analysis in Creole goats is still lacking.

This study is the first attempt towards a comprehensive genetic characterization of American Creole goats. Our findings indicate that the observed levels of genetic diversity are not high when compared with breeds from other parts of the world, possibly as a consequence of Creole goats being farther away from the domestication center and representing the last stages of dispersal and evolution but also due to founder effects related to the colonization process

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

underlying their origin. In general, Creole breeds represent well-differentiated entities and require measures aimed at their conservation and sustainable utilization. Dilution from admixture with commercial transboundary breeds appears to be negligible in most Creole breeds, but the majority of them displays significant levels of inbreeding. The genetic relationships among Creole breeds and the geographical patterns of genetic structure observed could reflect the effects of the alleged routes of expansion throughout the American continent during the colonization process, but a comprehensive study that includes goat breeds of Iberian and African origin is necessary to clarify the influence that breeds from these regions might have had in the development of Creole goats of the Americas.

Acknowledgements

The authors declare no conflict of interest. The authors wish to express thanks to breeder associations and research groups who kindly provided biological samples used in this study. Members of the CYTED XII-H and CONBIAND networks are thanked for valuable cooperation over the years. This study was partially supported by Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Spain (project references: RZ2004-00037-00-00, RZ2006- 00005-C02-00, RZ2007-00005-C02-01 and RZ2007-00005-C02-02), and by the Brazilian Government through CNPq and CAPES. CG was supported by a contract grant from the Fundação para a Ciência e a Tecnologia (IF/00866/2014), Portugal.

References

Ağaoğlu O.K. & Ertuğrul O. (2012) Assessment of genetic diversity, genetic relationship and bottleneck using microsatellites in some native Turkish goat breeds. *Small Ruminant Research* **105**, 53–60.

Agha S.H., Pilla F., Galal S., Shaat I., D'Andrea M., Reale S., Abdelsalam A.Z.A. & Li M.H. (2008) Genetic diversity in Egyptian and Italian goat breeds measured with microsatellite polymorphism. *Journal of Animal Breeding and Genetics* **125**, 194–200.

Ajmone-Marsan P., Colli L., Han J.L. *et al.* (2014) The characterization of goat genetic diversity: towards a genomic approach. *Small Ruminant Research* **121**, 58–72.

Aranguren-Méndez J., Portillo-Ríos M., Rincón X., Martínez A., Dickson L. & D'Aubeterre R. (2013) Diversidad genética en la cabra criolla venezolana mediante análisis con microsatelites. *Revista Cientifica FCV-LUZ* **XXIII**, 238–44.

Bandieri S., Favaro O. & Morinelli M. (1993) *Historia de Neuquén*. Ed. Plus Ultra, Buenos Aires.

Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. (1996–2004) GENETIX: *Logiciel sous Windows TM pour la Génétique des Populations*. Laboratoire Génome, Populations, Interactions, CNRS, Université de Montpellier II, Montpellier, France.

Boettcher P.J., Hoffmann I., Baumung R. *et al.* (2014) Genetic resources and genomics for adaptation of livestock to climate change. *Frontiers in Genetics* **5**, 461.

Bruford M.W., Ginja C., Hoffmann I. *et al.* (2015) Prospects and challenges for the conservation of farm animal genomic resources, 2015–2025. *Frontiers in Genetics* **6**, 314.

Bruno-de-Sousa C., Martinez A.M., Ginja C., Santos-Silva F., Carolino M.I., Delgado J.V. & Gama L.T. (2011) Genetic diversity and population structure in Portuguese goat breeds. *Livestock Science* **135**, 131–9.

Burren A., Neuditschko M., Signer-Hasler H., Frischknecht M., Reber I., Menzi F., Drogemuller C. & Flury C. (2016) Genetic diversity analyses reveal first insights into breed-specific selection signatures within Swiss goat breeds. *Animal Genetics* **47**, 727–39.

Calvo S.J., Gonzáles M.I., Ángel P.A., Cerón-Muñoz M.F. & Cardona-Cadavid H. (2012) Genetic evaluation of the goat population of Antioquia, using microsatellite markers. *Livestock Research for Rural Development* **24**; Article #82. Retrieved December 18, 2016, from http://www.lrrd.org/lrrd24/5/gonz24082.htm

Cañon J., García D., García-Atance M.A., Obexer-Ruff G., Lenstra J.A., Ajmone-Marsan P. & Dunner S. (2006) Geographical partitioning of goat diversity in Europe and the Middle East. *Animal Genetics* **37**, 327–34.

Chacón E., La O.M., Velásquez F.J., Pérez E., Delgado J.V., Cos Y., Fonseca Y. & Martínez A. (2012) Validation of the racial standards of the Cuban Creole goat for its international registration. *REDVET - Revista Electrónica de Veterinaria* **13**, 1–8.

Chapuis M.P. & Estoup A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**, 621–31.

Delgado J.V., Martinez A.M., Acosta A. *et al.* (2012) Genetic characterization of Latin-American Creole cattle using microsatellite markers. *Animal Genetics* **43**, 2–10.

Dempster A.P., Laird N.M. & Rubin D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **39**, 1–38.

Dixit S.P., Verma N.K., Aggarwal R.A.K., Vyas M.K., Rana J. & Sharma A. (2012) Genetic diversity and relationship among Indian goat breeds based on microsatellite markers. *Small Ruminant Research* **105**, 38–45.

Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–20.

Falush D., Stephens M. & Pritchard J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–87.

FAO (2007) *The State of the World's Animal Genetic Resources for Food and Agriculture* (Ed. by B. Rischkowsky & D. Pilling), p. 512. Food and Agriculture Organization of the United Nations, Rome, Italy.

Fatima S., Bhonga C.D., Ranka D.N. & Joshi C.G. (2008) Genetic variability and bottleneck studies in Zalawadi, Gohilwadi and Surti goat breeds of Gujarat (India) using microsatellites. *Small Ruminant Research* **77**, 58–64.

Ginja C., Penedo M.C., Melucci L., Quiroz J., Martinez Lopez O.R., Revidatti M.A., Martinez-Martinez A., Delgado J.V. & Gama L.T. (2010) Origins and genetic diversity of New World Creole cattle: inferences from mitochondrial and Y chromosome polymorphisms. *Animal Genetics* **41**, 128–41.

Ginja C., Gama L.T., Cortes O. *et al.* (2013) Analysis of conservation priorities of Iberoamerican cattle based on autosomal microsatellite markers. *Genetics Selection Evolution* **45**, 35.

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Goudet J. (1999) PCAGEN, *Vers. 1.2*. Institute of Ecology, University of Lausanne, Lausanne, Switzerland.

Goudet J. (2001) FSTAT: a program to estimate and test gene diversities and fixation indices. *Journal of Heredity* **86**, 485–6.

Groeneveld L.F., Lenstra J.A., Eding H. *et al.* (2010) Genetic diversity in farm animals—a review. *Animal Genetics* **41**(Suppl. 1), 6–31.

Guo S.W. & Thompson E.A. (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* **48**, 361–72.

Hartl D.L. & Clark A.G. (1997) *Principles of Population Genetics*. Sinauer Associates, Sunderland, MA, USA.

Hoffmann I. (2010) Climate change and the characterization, breeding and conservation of animal genetic resources. *Animal Genetics* **41**, 32–46.

Hoffmann I., Marsan P.A., Barker J.S.F., Cothran E.G., Hanotte O., Lenstra J.A., Milan D., Weigend S. & Simianer H. (2004) *New MoDAD Marker Sets to be Used in Diversity Studies for the Major Farm Animal Species: Recommendations of a Joint ISAG/FAO Working Group*. FAO, Rome.

Huson D.H. & Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254–67.

Iamartino D., Bruzzone A., Lanza A., Blasi M. & Pilla F. (2005) Genetic diversity of southern Italian goat populations assessed by microsatellite markers. *Small Ruminant Research* **57**, 249–55.

Kim S.C., Kang S.H., Choi E.Y., Hong Y.H., Bok J.D., Kim J.Y., Lee S.S., Choi Y.J., Choi I.S. & Cho K.K. (2016) Cloning and characterization of an endoglucanase gene from *Actinomyces* sp. Korean native goat 40. *Asian-Australas Journal of Animal Science* **29**, 126–33.

Langella O. (1999–2002) POPULATIONS *1.2.28*. CNRS UPR9034, France.

Lenstra J.A., Groeneveld L.F., Eding H. *et al.* (2012) Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. *Animal Genetics* **43**, 483–502.

Li M., Zhao S., Bian C. *et al.* (2002) Genetic relationships among twelve Chinese indigenous goat populations based on microsatellite analysis. *Genetics Selection Evolution* **34**, 729–44.

Mahmoudi B., Panahi B., Mohammadi S.A., Daliri M. & Babayev M.S. (2014) Microsatellite based phylogeny and bottleneck studies of Iranian indigenous goat populations. *Animal Biotechnology* **25**, 210–22.

Martinez A.M., Gama L.T., Delgado J.V. *et al.* (2015) The Southwestern fringe of Europe as an important reservoir of caprine biodiversity. *Genetics Selection Evolution* **47**, 86.

Menezes M.P., Martínez A.M., Ribeiro M.N., Pimenta-Filho E.C. & Bermejo J.V.D. (2006) Caracterização genética de raças caprinas nativas Brasileiras utilizando 27 marcadores microssatélites. *Revista Brasileira de Zootecnia* **35**, 1336–41.

Missohou A., Poutya M.R., Nenonene A., Dayo G.K., Ayssiwede S.B., Talaki E., Issa Y. & Fane A. (2011) Genetic diversity and differentiation in nine West African local goat breeds assessed via microsatellite polymorphism. *Small Ruminant Research* **99**, 20–4.

Naderi S., Rezaei H.R., Pompanon F. *et al.* (2008) The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17659–64.

Nei M., Tajima F. & Tateno Y. (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153–70.

Nicoloso L., Bomba L., Colli L. *et al.* (2015) Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genetics Selection Evolution* **47**, 62.

Oliveira J.C.V., Ribeiro M.N., Rocha L.L., Gomes-Filho M.A., Delgado J.V., Martinez A.M., Menezes M.P.C., Bettencourt C.M. & Gama L.T. (2010) Genetic relationships between two homologous goat breeds from Portugal and Brazil assessed by microsatellite markers. *Small Ruminant Research* **93**, 79–87.

Peakall R. & Smouse P.E. (2006) GENALEX 6: genetic analysis in EXCEL: population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–95.

Primo A.T. (1992) El ganado bovino Iberico en las Americas: 500 anos despues. *Archivos Zootecnia* **41**, 421–32.

Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–59.

Raymond M. & Rousset F. (2003) GENEPOP *(version 3.4). update of Raymond, M. & Rousset, F. (1995)* GENEPOP*: population genetics software for exact tests and ecumenicism*. Retrieved from http://genepop.curtin.edu.au/

Revidatti M.A., Bermejo J.V.D., Gama L.T., Periati V.L., Ginja C., Alvarez L.A., Vega-Pla J.L., Martinez A.M. & Consortium B. (2014) Genetic characterization of local Criollo pig breeds from the Americas using microsatellite markers. *Journal of Animal Science* **92**, 4823–32.

Ribeiro M.N., Bruno-de-Sousa C., Martinez-Martinez A., Ginja C., Menezes M.P.C., Pimenta E.C., Delgado J.V. & Gama L.T. (2012) Drift across the Atlantic: genetic differentiation and population structure in Brazilian and Portuguese native goat breeds. *Journal of Animal Breeding and Genetics* **129**, 79–87.

Rodero A., Delgado J.V. & Rodero E. (1992) Primitive Andalusian livestock and their implications in the discovery of America. *Archivos de Zootecnia* **41**, 383–400.

Serrano M., Calvo J.H., Martinez M., Marcos-Carcavilla A., Cuevas J., Gonzalez C., Jurado J.J. & de Tejada P.D. (2009) Microsatellite based genetic diversity and population structure of the endangered Spanish Guadarrama goat breed. *BMC Genetics* **10**, 61.

Sulabda N., Susari N.N.W., Heryani N.L.G.S. & Puja I.K. (2012) Genetic diversity of Gembrong Goat based on DNA microsatellite markers. *Indonesian Journal of Animal and Veterinary Science* **17**, 25–35.

Talenti A., Nicolazzi E.L., Chessa S. *et al.* (2016) A method for single nucleotide polymorphism selection for parentage assessment in goats. *Journal of Dairy Science* **99**, 3646–53.

Tosser-Klopp G., Bardou P., Bouchez O. *et al.* (2014) Design and characterization of a 52K SNP chip for goats. *PLoS One* **9**, e86227.

Traoré A., Álvarez I., Tamboura H.H. *et al.* (2009) Genetic characterisation of Burkina Faso goats using microsatellite polymorphism. *Livestock Science* **123**, 322–8.

Villalobos Cortés A.I., Martínez A.M., Vega-Pla J.L. & Delgado J.V. (2011) Genetic structure and bottleneck of the Guaymi bovine population by means of microsatellites. *Archivos de Zootecnia* **60**, 767–75.

Visser C., Lashmar S.F., Van Marle-Koster E., Poli M.A. & Allain D. (2016) Genetic diversity and population structure in South African, French and Argentinian Angora Goats from genome-wide SNP data. *PLoS ONE* **11**, e0154353.

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Walsh P.S., Metzger D.A. & Higuchi R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* **10**, 506–13.

Weir B.S. & Cockerham C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–70.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1 Modal distribution of ΔK values following Evanno *et al.* (2005) estimated for $K = 2$ to $K = 30$ for 24 Creole goat populations and three commercial transboundary goat breeds.

Figure S2 Results of the substructure analysis of six Creole goat breeds from Brazil and the commercial transboundary Alpine goat breed.

Figure S3 Results of the substructure analysis of six Creole goat breeds from Argentina and the commercial transboundary Alpine goat breed.

Table S1 Total number of alleles (TNA), mean number of alleles (MNA), effective number of alleles (n_e), allelic richness (R_t), populations showing private alleles (PA), observed (H_o) and expected (H_e) heterozygosities, Weir and Cockerham F Statistics (F_{is}, F_{it} and F_{st}) and number of populations showing heterozygote deficit (HWEd) estimated for each STR locus in 24 Creole and 3 commercial transboundary goat breeds.

Table S2 Pairwise-population F_{st} (above diagonal) and D_A (below diagonal) values.

Table S3 Average F_{st} estimates and D_A genetic distances for breed groups.

© 2017 Stichting International Foundation for Animal Genetics, doi: 10.1111/age.12529

Figure S1 Modal distribution of Δ*K* values following Evanno *et al.* 2005 estimated for *K* = 2 to *K* = 30 for 24 Creole goat populations and three commercial transboundary goat breeds.

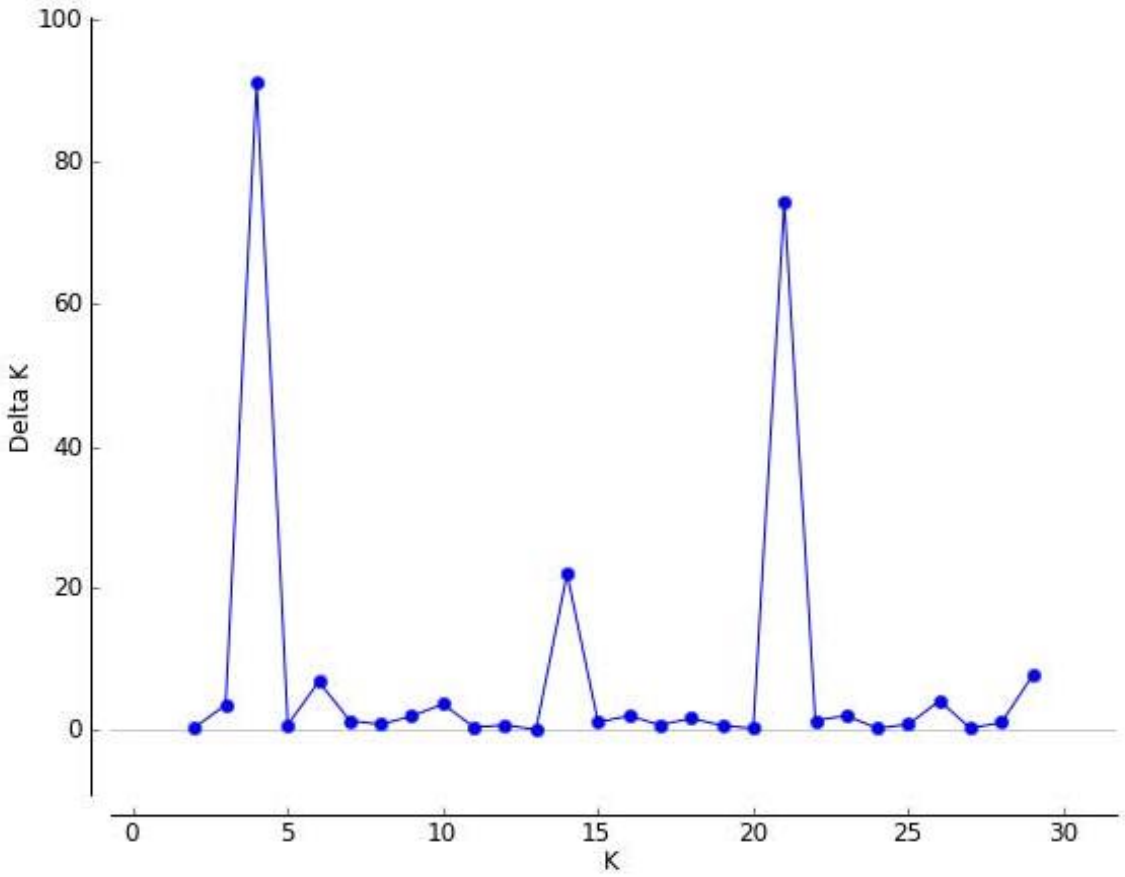
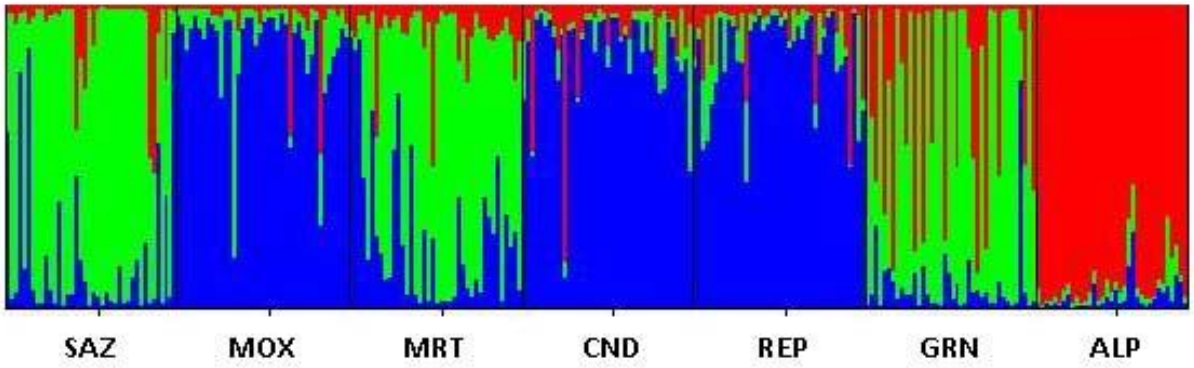
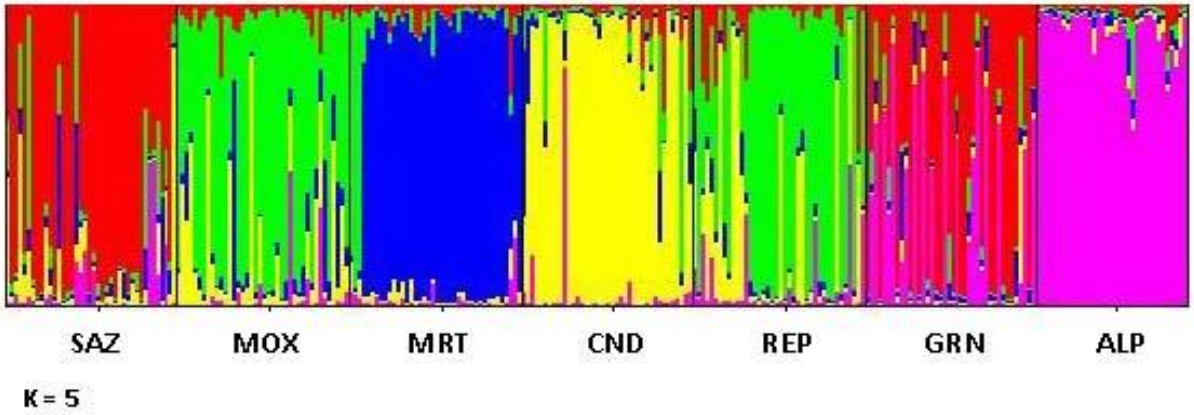


Figure S2 Results of the substructure analysis of six Creole goat breeds from Brazil and the commercial transboundary Alpine goat breed. Graphical representation of individual genotype membership coefficients (*q*) in each cluster obtained with STRUCTURE software without including prior information on sample origin for $K = 3$ and $K = 5$, and of the modal distribution of ΔK values following Evanno *et al.* 2005 for $K = 2$ to $K = 10$.





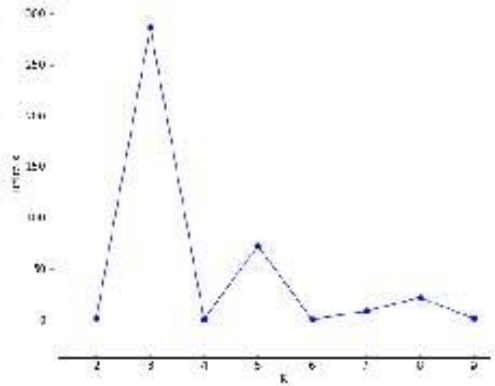


Figure S3 Results of the substructure analysis of six Creole goat breeds from Argentina and the commercial transboundary Alpine goat breed. Graphical representation of individual genotype membership coefficients (*q*) in each cluster obtained with STRUCTURE software without including prior information on sample origin for $K = 3$ and $K = 4$, and of the modal distribution of ΔK values following Evanno *et al.* 2005 for $K = 2$ to $K = 10$.

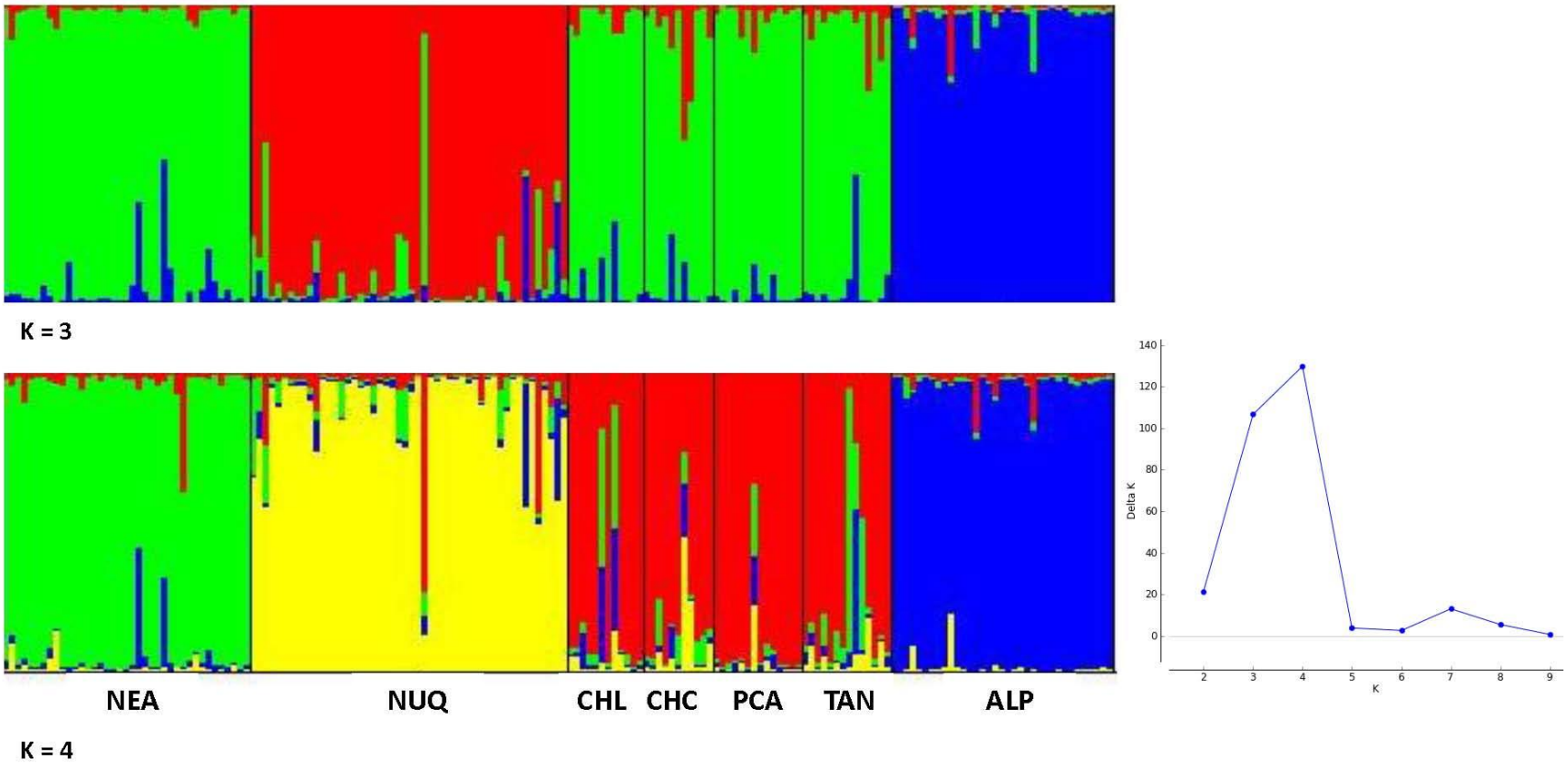


Table S1. Total number of alleles (TNA), mean number of alleles (MNA), effective number of alleles (n_e), allelic richness (R_t), populations showing private alleles (PA), observed (H_o) and expected (H_e) heterozygosities, Weir and Cockerham *F* Statistics (F_{is}, F_{it}, and F_{st}), number of populations showing heterozygote deficit (HWEd) estimated for each STR locus in 24 Creole and ŰKŴŜŜ commercial goat breeds.

Marker	TNA	MNA	n_e	R_t	PA	H_o	H_e	F_{it}	F_{st}	F_{is}	HWEd
BM1329	11	6.04	3.68	3.78	ALP	0.700	0.721	0.139	0.105	0.038	2
BM6506	10	6.07	3.26	3.89		0.530	0.643	0.341	0.217	0.158	10
BM6526	19	9.52	5.08	4.69		0.746	0.793	0.191	0.124	0.076	7
BM8125	9	5.44	3.35	3.59		0.680	0.665	0.138	0.148	-0.012	1
CRSM60	11	5.74	3.63	3.90		0.650	0.707	0.179	0.133	0.053	6
CSRD247	9	6.63	3.80	4.13		0.693	0.714	0.195	0.153	0.050	4
ETH010	6	3.63	2.27	2.86		0.493	0.537	0.224	0.174	0.061	3
ETH225	7	2.67	1.54	1.99	**ECU**; NUQ	0.273	0.319	0.305	0.128	0.203	6
HAUT27	11	5.11	2.67	3.14	MRT; PGY	0.583	0.605	0.128	0.090	0.042	5
ILSTS011	10	5.41	3.06	3.59		0.588	0.656	0.273	0.170	0.125	7
INRA063	7	4.41	2.50	2.76	PER	0.515	0.601	0.188	0.067	0.129	6
MAF065	14	7.56	4.59	4.30	**MYO**	0.658	0.763	0.255	0.124	0.150	11
MAF209	3	2.22	1.41	1.79		0.224	0.244	0.239	0.149	0.104	1
McM527	11	5.63	3.11	3.55	GAG; SAAN	0.592	0.651	0.217	0.146	0.082	9
MM12	18	8.52	4.52	4.52	**CHC**	0.760	0.768	0.129	0.120	0.010	2
OarFCB011	17	7.63	4.26	4.15	SAZ; PGY	0.679	0.733	0.191	0.145	0.055	7
OarFCB048	13	6.48	3.69	3.77		0.659	0.701	0.167	0.111	0.063	6
OarFCB304	21	8.74	3.82	4.26	VEN; PGY	0.649	0.724	0.234	0.132	0.117	9
SPS115	6	2.89	1.98	2.25	MOX; SAAN	0.419	0.495	0.275	0.088	0.204	12
SRCRSP08	12	6.59	3.72	4.03		0.649	0.715	0.228	0.147	0.095	10
TGLA122	11	6.63	3.87	4.04		0.668	0.715	0.194	0.137	0.067	5

PA in bold: frequency > 5%. HWEd for $P < 0.05$.

Table S2. Pairwise-population F_{st} (above diagonal) and D_A (below diagonal) values. Maximum and minimum values are shown in bold.

	Breeds	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	SPA		0.057	0.200	0.097	0.054	0.062	0.059	0.207	0.055	0.153	0.122	0.112	0.099	0.086	0.100	0.046	0.091	0.032	0.047	0.061	0.054	0.056	0.070	0.054	0.054	0.055	0.083
2	MYO	0.128		0.275	0.133	0.091	0.092	0.118	0.267	0.086	0.207	0.146	0.139	0.141	0.120	0.153	0.083	0.132	0.070	0.088	0.104	0.102	0.097	0.092	0.086	0.095	0.080	0.125
3	SCL	0.297	0.366		0.296	0.244	0.263	0.297	**0.391**	0.239	0.361	0.316	0.331	0.276	0.295	0.310	0.225	0.269	0.196	0.230	0.272	0.293	0.291	0.306	0.291	0.228	0.249	0.266
4	CUB	0.188	0.237	0.355		0.151	0.127	0.171	0.276	0.091	0.250	0.187	0.191	0.154	0.149	0.168	0.122	0.125	0.098	0.107	0.124	0.108	0.132	0.151	0.132	0.129	0.124	0.179
5	COL	0.139	0.198	0.316	0.233		0.101	0.095	0.235	0.098	0.196	0.148	0.147	0.120	0.135	0.147	0.065	0.123	0.051	0.081	0.087	0.077	0.088	0.111	0.086	0.095	0.075	0.087
6	VEN	0.152	0.187	0.337	0.186	0.186		0.093	0.239	0.060	0.228	0.194	0.184	0.169	0.160	0.177	0.121	0.104	0.077	0.076	0.109	0.103	0.096	0.115	0.084	0.094	0.099	0.155
7	ECU	0.172	0.249	0.368	0.305	0.215	0.201		0.276	0.087	0.248	0.208	0.206	0.164	0.158	0.196	0.105	0.140	0.069	0.080	0.088	0.106	0.098	0.129	0.099	0.112	0.109	0.108
8	GAG	0.361	0.414	0.480	0.400	0.368	0.314	0.377		0.209	0.377	0.302	0.327	0.273	0.275	0.319	0.224	0.216	0.183	0.206	0.213	0.232	0.233	0.266	0.239	0.221	0.231	0.299
9	PER	0.126	0.171	0.303	0.163	0.166	0.110	0.200	0.320		0.208	0.181	0.164	0.145	0.147	0.144	0.090	0.072	0.064	0.047	0.066	0.067	0.067	0.076	0.064	0.078	0.085	0.137
10	SAZ	0.226	0.298	0.421	0.307	0.256	0.305	0.350	**0.494**	0.266		0.092	0.089	0.102	0.106	0.040	0.128	0.216	0.154	0.186	0.217	0.188	0.196	0.212	0.181	0.200	0.154	0.236
11	MOX	0.225	0.263	0.392	0.267	0.244	0.303	0.344	0.449	0.272	0.118		0.069	0.038	0.046	0.080	0.084	0.193	0.126	0.153	0.175	0.159	0.170	0.194	0.147	0.155	0.131	0.198
12	MRT	0.196	0.253	0.429	0.251	0.245	0.272	0.318	0.449	0.241	0.133	0.108		0.095	0.089	0.068	0.088	0.176	0.130	0.149	0.157	0.144	0.151	0.147	0.121	0.165	0.109	0.197
13	CND	0.184	0.267	0.358	0.218	0.197	0.265	0.299	0.409	0.219	0.127	0.095	0.139		0.050	0.086	0.066	0.172	0.104	0.122	0.146	0.128	0.138	0.160	0.128	0.131	0.102	0.164
14	REP	0.167	0.227	0.379	0.210	0.206	0.254	0.286	0.418	0.224	0.127	**0.074**	0.119	0.093		0.088	0.069	0.171	0.098	0.126	0.151	0.132	0.144	0.163	0.130	0.118	0.106	0.160
15	GRN	0.172	0.243	0.373	0.230	0.228	0.270	0.332	0.470	0.208	0.080	0.127	0.121	0.132	0.129		0.084	0.158	0.114	0.133	0.156	0.127	0.137	0.149	0.131	0.144	0.104	0.184
16	SRD	0.118	0.203	0.308	0.211	0.164	0.224	0.239	0.386	0.177	0.178	0.147	0.154	0.134	0.108	0.140		0.116	0.053	0.081	0.101	0.072	0.077	0.089	0.082	0.085	0.072	0.085
17	BOL	0.217	0.272	0.347	0.206	0.252	0.200	0.280	0.329	0.138	0.316	0.325	0.282	0.291	0.296	0.258	0.253		0.093	0.056	0.117	0.084	0.091	0.089	0.078	0.113	0.103	0.185
18	PGY	0.087	0.166	0.299	0.189	0.146	0.174	0.204	0.354	0.123	0.260	0.247	0.247	0.208	0.186	0.202	0.141	0.205		0.041	0.078	0.054	0.059	0.081	0.051	0.073	0.059	0.064
19	NEA	0.130	0.186	0.288	0.179	0.171	0.157	0.212	0.342	0.092	0.271	0.262	0.250	0.205	0.225	0.206	0.176	0.130	0.102		0.071	0.044	0.043	0.064	0.040	0.087	0.068	0.126
20	NUQ	0.140	0.193	0.349	0.206	0.176	0.199	0.235	0.329	0.144	0.268	0.252	0.233	0.213	0.219	0.219	0.194	0.226	0.168	0.149		0.061	0.056	0.088	0.056	0.093	0.084	0.133
21	CHL	0.152	0.224	0.379	0.236	0.209	0.219	0.273	0.382	0.162	0.284	0.301	0.259	0.241	0.250	0.222	0.198	0.228	0.174	0.141	0.145		***-0.008***	*0.017*	*0.020*	0.092	0.075	0.136
22	CHC	0.159	0.205	0.369	0.266	0.221	0.221	0.255	0.368	0.164	0.298	0.312	0.272	0.258	0.273	0.250	0.211	0.241	0.178	0.150	0.142	0.096		0.021	*0.021*	0.105	0.069	0.131
23	PCA	0.180	0.224	0.401	0.308	0.253	0.247	0.288	0.387	0.188	0.341	0.370	0.282	0.294	0.314	0.288	0.252	0.252	0.231	0.193	0.185	0.115	0.116		*0.027*	0.119	0.082	0.161
24	TAN	0.171	0.218	0.379	0.277	0.222	0.194	0.254	0.353	0.164	0.293	0.281	0.247	0.249	0.259	0.243	0.218	0.206	0.189	0.144	0.147	0.146	0.137	0.147		0.098	0.071	0.128
25	SAAN	0.139	0.181	0.298	0.231	0.190	0.184	0.255	0.338	0.155	0.252	0.238	0.248	0.212	0.198	0.183	0.178	0.223	0.163	0.176	0.175	0.207	0.226	0.230	0.201		0.067	0.133
26	ALP	0.145	0.189	0.345	0.231	0.194	0.221	0.261	0.380	0.192	0.225	0.235	0.196	0.198	0.191	0.165	0.180	0.238	0.157	0.168	0.183	0.185	0.200	0.212	0.194	0.146		0.130
27	ANG	0.138	0.206	0.336	0.270	0.174	0.229	0.203	0.480	0.213	0.325	0.302	0.296	0.264	0.225	0.274	0.156	0.332	0.147	0.217	0.242	0.263	0.242	0.304	0.272	0.237	0.240	

F_{st} values were all significant for $P < 0.05$ except for the following pairwise-population values shown in italic: CHL/CHC, CHL/PCA, CHL/TAN, CHC/TAN, and PCA/TAN.

Table S3. Average F_{st} estimates and D_A genetic distances for breed groups.

Breed group	F_{st} ± SD	D_A ± SD
Brazil	0.079 ± 0.022	0.123 ± 0.024
Brazil & Transboundary	0.111 ± 0.046	0.177 ± 0.063
Argentina	0.041 ± 0.025	0.144 ± 0.024
Argentina & Transboundary	0.078 ± 0.041	0.185 ± 0.048
Creole (no Island goats)	0.111 ± 0.048	0.210 ± 0.061
Creole (no Island goats) & Transboundary	0.112 ± 0.047	0.210 ± 0.059
Creole	0.138 ± 0.075	0.237 ± 0.085
Creole & Transboundary	0.134 ± 0.071	0.234 ± 0.080

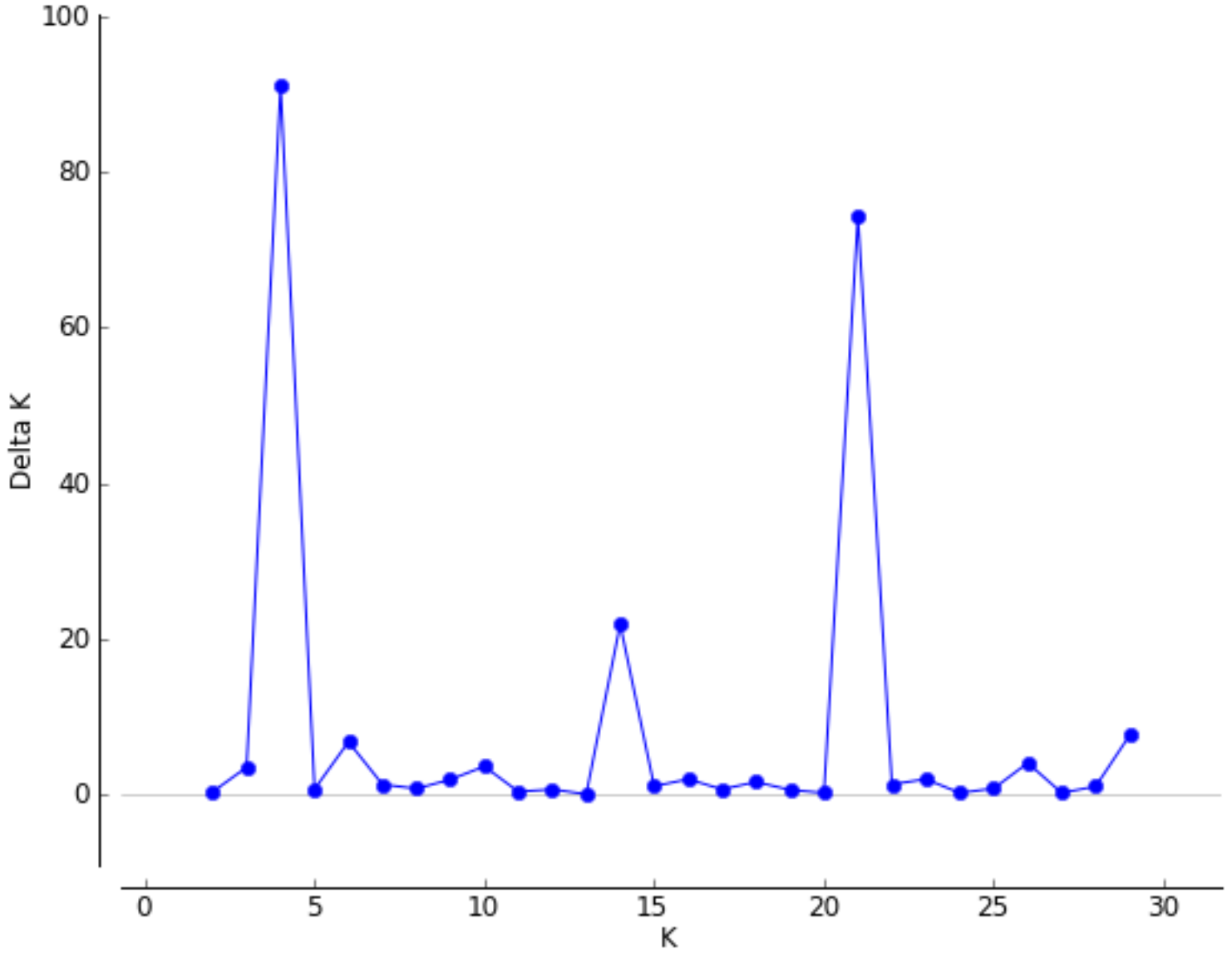
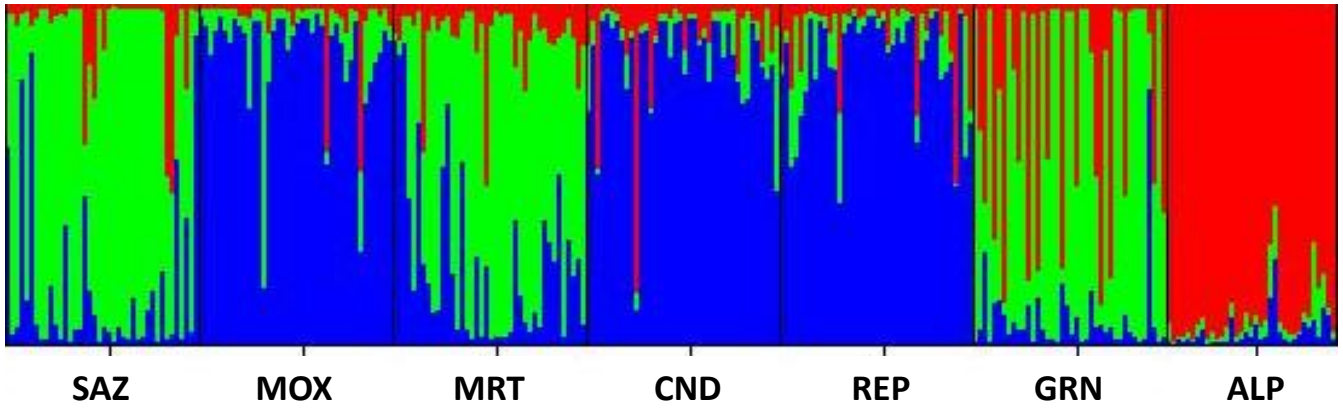
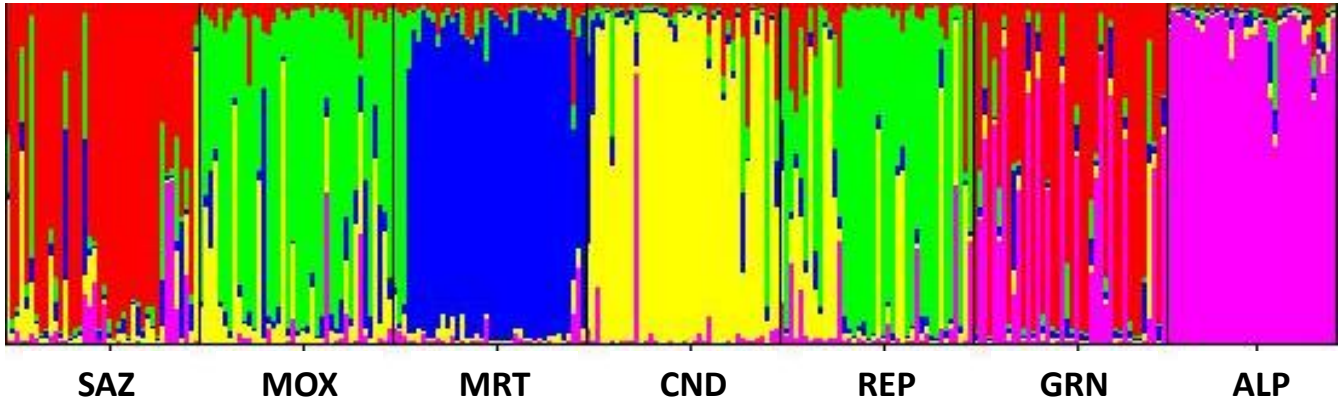


Figure S1. Graphic showing the modal distribution of *ΔK* values following Evanno *et al.* 2005 estimated for *K* = 2 to *K* = 30 for 24 Creole goat populations and three commercial transboundary goat breeds.



K = 3



K = 5

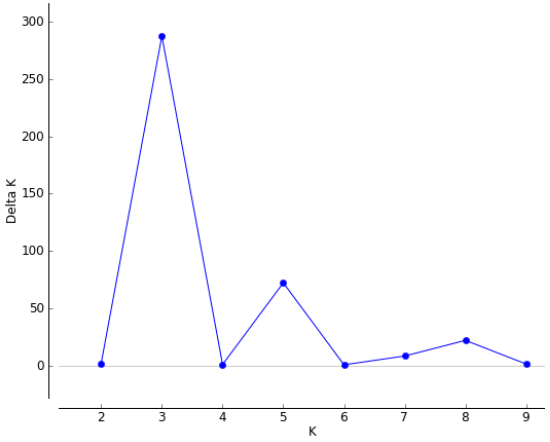
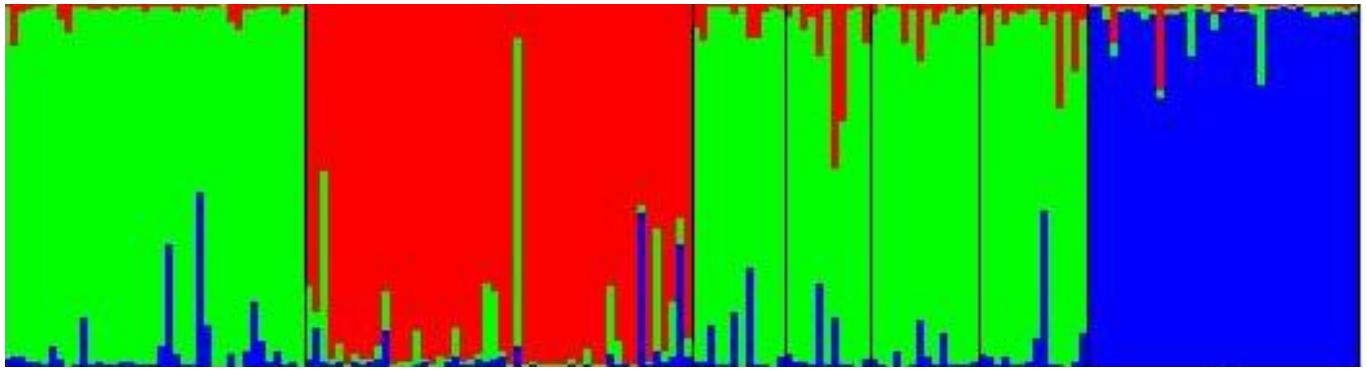
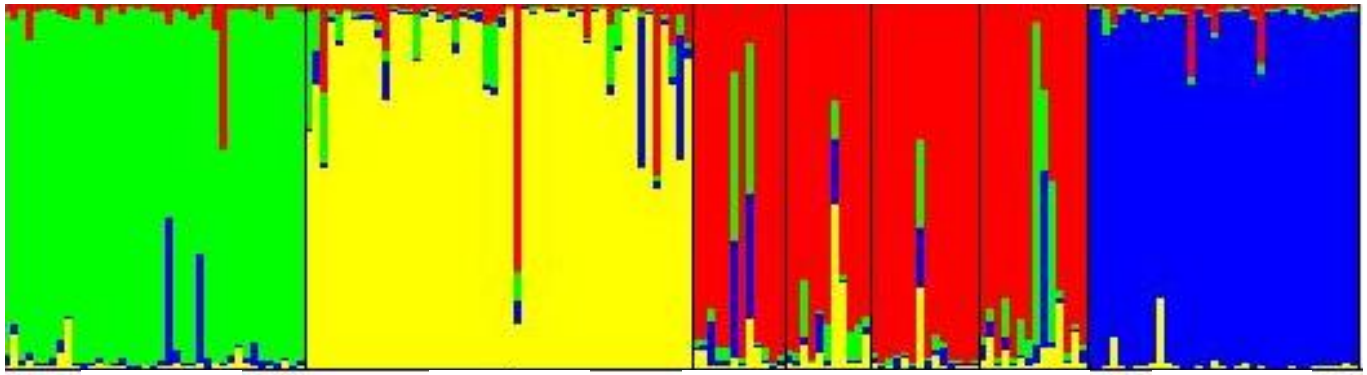


Figure S2. Results of the substructure analysis of six Creole goat breeds from Brazil and the commercial transboundary Alpine goat breed. Graphical representation of individual genotype membership coefficients (*q*) in each cluster obtained with Structure software without including prior information on sample origin for *K* = 3 and *K* = 5, and of the modal distribution of *ΔK* values following Evanno *et al.* 2005 for *K* = 2 to *K* = 10.



K = 3



K = 4

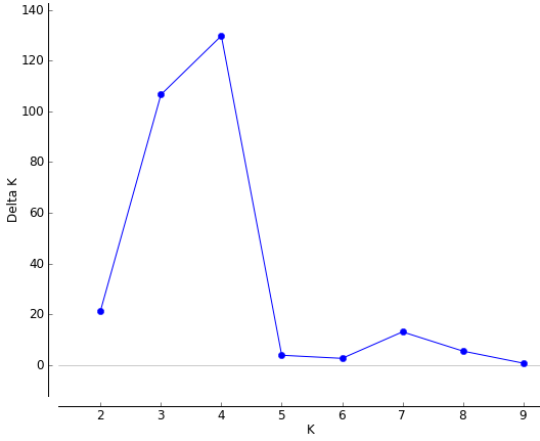


Figure S3. Results of the substructure analysis of six Creole goat breeds from Argentina and the commercial transboundary Alpine goat breed. Graphical representation of individual genotype membership coefficients (*q*) in each cluster obtained with Structure software without including prior information on sample origin for *K* = 3 and *K* = 4, and of the modal distribution of *ΔK* values following Evanno *et al.* 2005 for *K* = 2 to *K* = 10.

Table S2. Pairwise-population F_{st} (above diagonal) and D_A (below diagonal) values. Maximum and minimum values are shown in bold.

	Breeds	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	SPA		0.057	0.200	0.097	0.054	0.062	0.059	0.207	0.055	0.153	0.122	0.112	0.099	0.086	0.100	0.046	0.091	0.032	0.047	0.061	0.054	0.056	0.070	0.054	0.054	0.055	0.083
2	MYO	0.128		0.275	0.133	0.091	0.092	0.118	0.267	0.086	0.207	0.146	0.139	0.141	0.120	0.153	0.083	0.132	0.070	0.088	0.104	0.102	0.097	0.092	0.086	0.095	0.080	0.125
3	SCL	0.297	0.366		0.296	0.244	0.263	0.297	**0.391**	0.239	0.361	0.316	0.331	0.276	0.295	0.310	0.225	0.269	0.196	0.230	0.272	0.293	0.291	0.306	0.291	0.228	0.249	0.266
4	CUB	0.188	0.237	0.355		0.151	0.127	0.171	0.276	0.091	0.250	0.187	0.191	0.154	0.149	0.168	0.122	0.125	0.098	0.107	0.124	0.108	0.132	0.151	0.132	0.129	0.124	0.179
5	COL	0.139	0.198	0.316	0.233		0.101	0.095	0.235	0.098	0.196	0.148	0.147	0.120	0.135	0.147	0.065	0.123	0.051	0.081	0.087	0.077	0.088	0.111	0.086	0.095	0.075	0.087
6	VEN	0.152	0.187	0.337	0.186	0.186		0.093	0.239	0.060	0.228	0.194	0.184	0.169	0.160	0.177	0.121	0.104	0.077	0.076	0.109	0.103	0.096	0.115	0.084	0.094	0.099	0.155
7	ECU	0.172	0.249	0.368	0.305	0.215	0.201		0.276	0.087	0.248	0.208	0.206	0.164	0.158	0.196	0.105	0.140	0.069	0.080	0.088	0.106	0.098	0.129	0.099	0.112	0.109	0.108
8	GAG	0.361	0.414	0.480	0.400	0.368	0.314	0.377		0.209	0.377	0.302	0.327	0.273	0.275	0.319	0.224	0.216	0.183	0.206	0.213	0.232	0.233	0.266	0.239	0.221	0.231	0.299
9	PER	0.126	0.171	0.303	0.163	0.166	0.110	0.200	0.320		0.208	0.181	0.164	0.145	0.147	0.144	0.090	0.072	0.064	0.047	0.066	0.067	0.067	0.076	0.064	0.078	0.085	0.137
10	SAZ	0.226	0.298	0.421	0.307	0.256	0.305	0.350	**0.494**	0.266		0.092	0.089	0.102	0.106	0.040	0.128	0.216	0.154	0.186	0.217	0.188	0.196	0.212	0.181	0.200	0.154	0.236
11	MOX	0.225	0.263	0.392	0.267	0.244	0.303	0.344	0.449	0.272	0.118		0.069	0.038	0.046	0.080	0.084	0.193	0.126	0.153	0.175	0.159	0.170	0.194	0.147	0.155	0.131	0.198
12	MRT	0.196	0.253	0.429	0.251	0.245	0.272	0.318	0.449	0.241	0.133	0.108		0.095	0.089	0.068	0.088	0.176	0.130	0.149	0.157	0.144	0.151	0.147	0.121	0.165	0.109	0.197
13	CND	0.184	0.267	0.358	0.218	0.197	0.265	0.299	0.409	0.219	0.127	0.095	0.139		0.050	0.086	0.066	0.172	0.104	0.122	0.146	0.128	0.138	0.160	0.128	0.131	0.102	0.164
14	REP	0.167	0.227	0.379	0.210	0.206	0.254	0.286	0.418	0.224	0.127	**0.074**	0.119	0.093		0.088	0.069	0.171	0.098	0.126	0.151	0.132	0.144	0.163	0.130	0.118	0.106	0.160
15	GRN	0.172	0.243	0.373	0.230	0.228	0.270	0.332	0.470	0.208	0.080	0.127	0.121	0.132	0.129		0.084	0.158	0.114	0.133	0.156	0.127	0.137	0.149	0.131	0.144	0.104	0.184
16	SRD	0.118	0.203	0.308	0.211	0.164	0.224	0.239	0.386	0.177	0.178	0.147	0.154	0.134	0.108	0.140		0.116	0.053	0.081	0.101	0.072	0.077	0.089	0.082	0.085	0.072	0.085
17	BOL	0.217	0.272	0.347	0.206	0.252	0.200	0.280	0.329	0.138	0.316	0.325	0.282	0.291	0.296	0.258	0.253		0.093	0.056	0.117	0.084	0.091	0.089	0.078	0.113	0.103	0.185
18	PGY	0.087	0.166	0.299	0.189	0.146	0.174	0.204	0.354	0.123	0.260	0.247	0.247	0.208	0.186	0.202	0.141	0.205		0.041	0.078	0.054	0.059	0.081	0.051	0.073	0.059	0.064
19	NEA	0.130	0.186	0.288	0.179	0.171	0.157	0.212	0.342	0.092	0.271	0.262	0.250	0.205	0.225	0.206	0.176	0.130	0.102		0.071	0.044	0.043	0.064	0.040	0.087	0.068	0.126
20	NUQ	0.140	0.193	0.349	0.206	0.176	0.199	0.235	0.329	0.144	0.268	0.252	0.233	0.213	0.219	0.219	0.194	0.226	0.168	0.149		0.061	0.056	0.088	0.056	0.093	0.084	0.133
21	CHL	0.152	0.224	0.379	0.236	0.209	0.219	0.273	0.382	0.162	0.284	0.301	0.259	0.241	0.250	0.222	0.198	0.228	0.174	0.141	0.145		***-0.008***	*0.017*	*0.020*	0.092	0.075	0.136
22	CHC	0.159	0.205	0.369	0.266	0.221	0.221	0.255	0.368	0.164	0.298	0.312	0.272	0.258	0.273	0.250	0.211	0.241	0.178	0.150	0.142	0.096		0.021	*0.021*	0.105	0.069	0.131
23	PCA	0.180	0.224	0.401	0.308	0.253	0.247	0.288	0.387	0.188	0.341	0.370	0.282	0.294	0.314	0.288	0.252	0.252	0.231	0.193	0.185	0.115	0.116		*0.027*	0.119	0.082	0.161
24	TAN	0.171	0.218	0.379	0.277	0.222	0.194	0.254	0.353	0.164	0.293	0.281	0.247	0.249	0.259	0.243	0.218	0.206	0.189	0.144	0.147	0.146	0.137	0.147		0.098	0.071	0.128
25	SAAN	0.139	0.181	0.298	0.231	0.190	0.184	0.255	0.338	0.155	0.252	0.238	0.248	0.212	0.198	0.183	0.178	0.223	0.163	0.176	0.175	0.207	0.226	0.230	0.201		0.067	0.133
26	ALP	0.145	0.189	0.345	0.231	0.194	0.221	0.261	0.380	0.192	0.225	0.235	0.196	0.198	0.191	0.165	0.180	0.238	0.157	0.168	0.183	0.185	0.200	0.212	0.194	0.146		0.130
27	ANG	0.138	0.206	0.336	0.270	0.174	0.229	0.203	0.480	0.213	0.325	0.302	0.296	0.264	0.225	0.274	0.156	0.332	0.147	0.217	0.242	0.263	0.242	0.304	0.272	0.237	0.240	

F_{st} values were all significant for $P < 0.05$ except for the following pairwise-population values shown in italic: CHL/CHC, CHL/PCA, CHL/TAN, CHC/TAN, and PCA/TAN.

Table S3. Average F_{st} estimates and D_A genetic distances for breed groups.

Breed group	F_{st} ± SD	D_A ± SD
Brazil	0.079 ± 0.022	0.123 ± 0.024
Brazil & Transboundary	0.111 ± 0.046	0.177 ± 0.063
Argentina	0.041 ± 0.025	0.144 ± 0.024
Argentina & Transboundary	0.078 ± 0.041	0.185 ± 0.048
Creole (no Island goats)	0.111 ± 0.048	0.210 ± 0.061
Creole (no Island goats) & Transboundary	0.112 ± 0.047	0.210 ± 0.059
Creole	0.138 ± 0.075	0.237 ± 0.085
Creole & Transboundary	0.134 ± 0.071	0.234 ± 0.080