Animal (2018), **12:10**, pp 2017–2026 © The Animal Consortium 2018
doi:10.1017/S1751731117003627

Dissection of ancestral genetic contributions to Creole goat populations

N. Sevane[1†], O. Cortés[1], L. T. Gama[2], A. Martínez[3,4], P. Zaragoza[5], M. Amills[6], D. O. Bedotti[7], C. Bruno de Sousa[8], J. Cañon[1], S. Dunner[1], C. Ginja[9], M. R. Lanari[10], V. Landi[3,4], P. Sponenberg[11], J. V. Delgado[3] and The BioGoat Consortium[a]

[1]Departamento de Producción Animal, Universidad Complutense de Madrid, Spain; [2]CIISA, Faculdade de Medicina Veterinária, Universidade de Lisboa, Portugal; [3]Departamento de Genética, Universidad de Córdoba, Córdoba, Spain; [4]Animal Breeding Consulting SL, Cordoba, Spain; [5]Laboratorio de Genética Bioquímica, Facultad de Veterinaria, Universidad de Zaragoza, Spain; [6]Department of Animal Genetics, Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Universitat Autònoma de Barcelona, Bellaterra, Spain; [7]INTA, EEA Anguil "Ing. Agr. Guillermo Covas", Argentina; [8]Centro de Ciências do Mar, Instituto de Higiene e Medicina Tropical (UPMM) – UNL, Universidade do Algarve, Portugal; [9]Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-InBIO), Universidade do Porto, Campus Agrário de Vairão, Portugal; [10]Area de Producción Animal, INTA EEA, Bariloche, Argentina; [11]Virginia-Maryland Regional College of Veterinary Medicine, Virginia Tech, Blacksburg, VA, USA

(Received 19 September 2017; Accepted 4 December 2017; First published online 8 January 2018)

Goats have played a key role as source of nourishment for humans in their expansion all over the world in long land and sea trips. This has guaranteed a place for this species in the important and rapid episode of livestock expansion triggered by Columbus' arrival in the Americas in the late 1400s. The aims of this study are to provide a comprehensive perspective on genetic diversity in American goat populations and to assess their origins and evolutionary trajectories. This was achieved by combining data from autosomal neutral genetic markers obtained in more than two thousand samples that encompass a wide range of Iberian, African and Creole goat breeds. In general, even though Creole populations differ clearly from each other, they lack a strong geographical pattern of differentiation, such that populations of different admixed ancestry share relatively close locations throughout the large geographical range included in this study. Important Iberian signatures were detected in most Creole populations studied, and many of them, particularly the Cuban Creole, also revealed an important contribution of African breeds. On the other hand, the Brazilian breeds showed a particular genetic structure and were clearly separated from the other Creole populations, with some influence from Cape Verde goats. These results provide a comprehensive characterisation of the present structure of goat genetic diversity, and a dissection of the Iberian and African influences that gave origin to different Creole caprine breeds, disentangling an important part of their evolutionary history. Creole breeds constitute an important reservoir of genetic diversity that justifies the development of appropriate management systems aimed at improving performance without loss of genomic diversity.

Keywords: *Capra hircus*, Iberia, America, microsatellites, genetic diversity

Implications

A better understanding of the genetic diversity and structure of Creole goat populations is crucial for disentangling their evolutionary history and define adequate conservation priorities to maintain the vast genetic legacy that we have inherited. The hardiness and adaptiveness that goats display in harsh environments, where other livestock species find it difficult to adapt, make the genetic reservoir of locally adapted Creole breeds a valuable source of diversity to maintain and improve caprine performance and welfare under expected climatic change models.

Introduction

The domestic goat (*Capra hircus*) was one of the earliest animals to be domesticated and remains a species of major importance in poorly developed rural areas. Given their small

---

[†] E-mail: nsevane@ucm.es

[a] List of the members of the BioGoat Consortium (full information in Supplementary Material): Tumininu Adebambo; Lenin Aguirre Riofrío; Luz Angela Alvarez Franco; Jose Atilio Aranguren; Mª Esperanza Camacho Vallejo; Juan Capote; María del Rosario Fresno; Mª Ines Carolino; Fátima Santos Silva; Marcos P. Carrera Menezes; Edilberto Chacón; Ahmed R.A.E Elbeltagy; Gabriel Fernández; Ainhoa Ferrando; Jordi Jordana; Angel Galarza; Angelika Stemmer; Daniel Martín Santana; Mayra Gómez Carpio; Mariano Gómez; Nilton Gómez Urviola; Arianna Manunza; Inmaculada Martín-Burriel; Clementina Rodellar; Roberto Martínez; Pere-Miquel Parés i Casanova; Edgar C. Pimenta-Filho; Agueda Pons Barro; Mª Antonia Revidatti; Maria Norma Ribeiro; Laura Leandro Rocha; Asmaa Mohammed Aly Abu Shady; Bartomeu Seguí; Jose Luis Vega Pla; Oriol Vidal.

size and high versatility, goats were extensively transported among continents, weakening their phylogeographic structure at a worldwide level, such that most genetic variation is currently found at the within-population level (Cañón *et al.*, 2006; Naderi *et al.*, 2007; Ginja *et al.*, 2018). Columbus' arrival in the Americas triggered a major change, with the expansion of European and African livestock species in the New World, starting from a limited number of animals brought around the beginning of the 16th century (Rodero *et al.*, 1992). The development of goat populations in the American continents has thus been shaped by founder effects, genetic drift, and natural and artificial selection, leading to a progressive divergence from their original founder populations.

The study of goat genetic diversity can bring insights into their evolutionary history, and can explain the present geographical distribution and genetic structure of this species. Although several studies have been carried out to characterise the genetic diversity of European, African, and South and Central American goats at the continental level (e.g. Cañón *et al.*, 2006; Naderi *et al.*, 2007; Ginja *et al.*, 2017), a comprehensive joint analysis of breeds from the various continents has not been accomplished so far. Some research on New World goats has relied on mitochondrial DNA sequences (Amills *et al.*, 2009), which reflect only the history of female-mediated gene flow and thus lack the power to track complex histories of introgression and admixture. Despite other studies with samples of Portuguese and Brazilian origin (Ribeiro *et al.*, 2012), the likely influence that the other Iberian breeds, as well as African populations, may have had on the current genetic structure of Creole goats has not been investigated so far.

The aims of this study are to provide a comprehensive assessment of the genetic diversity, origins and evolutionary trajectories of American goat populations by combining data from autosomal neutral genetic markers collected in >2000 samples that include a wide range of Iberian, African and Creole caprine breeds. The populations sampled cover a large geographical distribution (Figure 1) and encompass the majority of the extant caprine breeds from the Americas and Iberia, and 13 African local breeds spanning from Egypt to South Africa. Using a set of 20 microsatellite markers, we show that some of the diversity displayed by Creole populations derives from admixture between Iberian and African lineages. On the other hand, Creole goat breeds reveal their unique identity, making them an important reservoir of genetic diversity for selection and conservation programmes.

Methods

Sampling and microsatellite analyses

Biological samples (blood or hair roots) were collected from a total of 2395 unrelated purebred goats registered in Herdbooks (whenever available) of 71 populations from diverse geographic origins and belonging to 69 recognised breeds. The whole sample included (Figure 1, Supplementary Table S1): 23 populations belonging to 21 local breeds from Spain, with duplicated samples representing the Blanca Celtibérica and Celtibérica, two well-differentiated varieties of the same breed, and Tenerife Norte and Tenerife Sur, which are ecotypes of the Tinerfeña breed (Martínez *et al.*, 2015); six breeds from Portugal; 13 breeds from Africa; two breeds from New Zealand; the Golden Guernsey breed from the United Kingdom, sampled in United States; three cosmopolitan breeds, that is, Saanen from Switzerland (sampled in Spain), Alpine from France and Anglo-Nubian from the United Kingdom (both sampled in Brazil); and 23 Creole breeds distributed throughout the Americas, from the United States to Argentina. The number of samples per breed ranged between 11 and 40, with an average of 34 individuals. Samples were collected by qualified veterinarians through their routine practice, in the framework of official programs aimed at identification, health control and parentage confirmation. Therefore, no ethical approval was required for sampling biological material. Genomic DNA was isolated as described by Walsh *et al.* (1991). A set of 20 microsatellite markers was selected (Supplementary Table S2) according to the recommendations of the Food and Agriculture Organization of the United Nations (FAO) and the International Society for Animal Genetics (FAO, 2011), and amplified in multiplex PCRs as described in Martínez *et al.* (2015).

Statistical analyses

The MICROSATELLITE TOOLKIT software (Park, 2001) was used to calculate allele frequencies for each locus, total number of alleles per locus (NA), observed (H_o) and expected (H_e) heterozygosities, and mean number of alleles per population. Weir and Cockerham (1984) estimates of *F*-statistics and allelic richness (Ar) over all loci per breed were estimated with FSTAT (Goudet, 2001). The mean *F*-statistics and their 95% confidence interval over loci after 1000 bootstraps were calculated using the GENETIX 4.05 software (Belkhir *et al.*, 2004). Compliance to Hardy–Weinberg (HW) equilibrium was analysed using Fisher's exact test with GENEPOP 1.2 (Raymond and Rousset, 1995).

Factorial correspondence analysis (FCA) using the function 'AFC 3D by populations' was performed with GENETIX. Genetic relationships among populations and breed groups were analysed using Reynolds genetic distances implemented in PHYLIP 3.69 (Felsenstein, 2004). The neighbour-net method, as implemented in SPLITSTREE software (Huson and Bryant, 2006), was used to compute a network based on Reynolds distances to graphically represent breed relationships and to depict evidence of admixture.

The model-based clustering software STRUCTURE 2.2 (Pritchard *et al.*, 2000) was used to infer genetic contributions from ancestral populations and population substructure in each goat breed with the admixture model and uncorrelated allele frequencies. For each value of the putative number of genetic clusters (K) ranging between 2 and 35, five independent runs were performed with Monte-Carlo Markov-Chain iterations ranging from 250 000 (50 000 burn-in length) for $K = 2$ to 10 to 500 000 (150 000 burn-in length) for $K = 11$ to 35, and the results were visualised with the DISTRUCT software (Rosenberg, 2004). The most probable

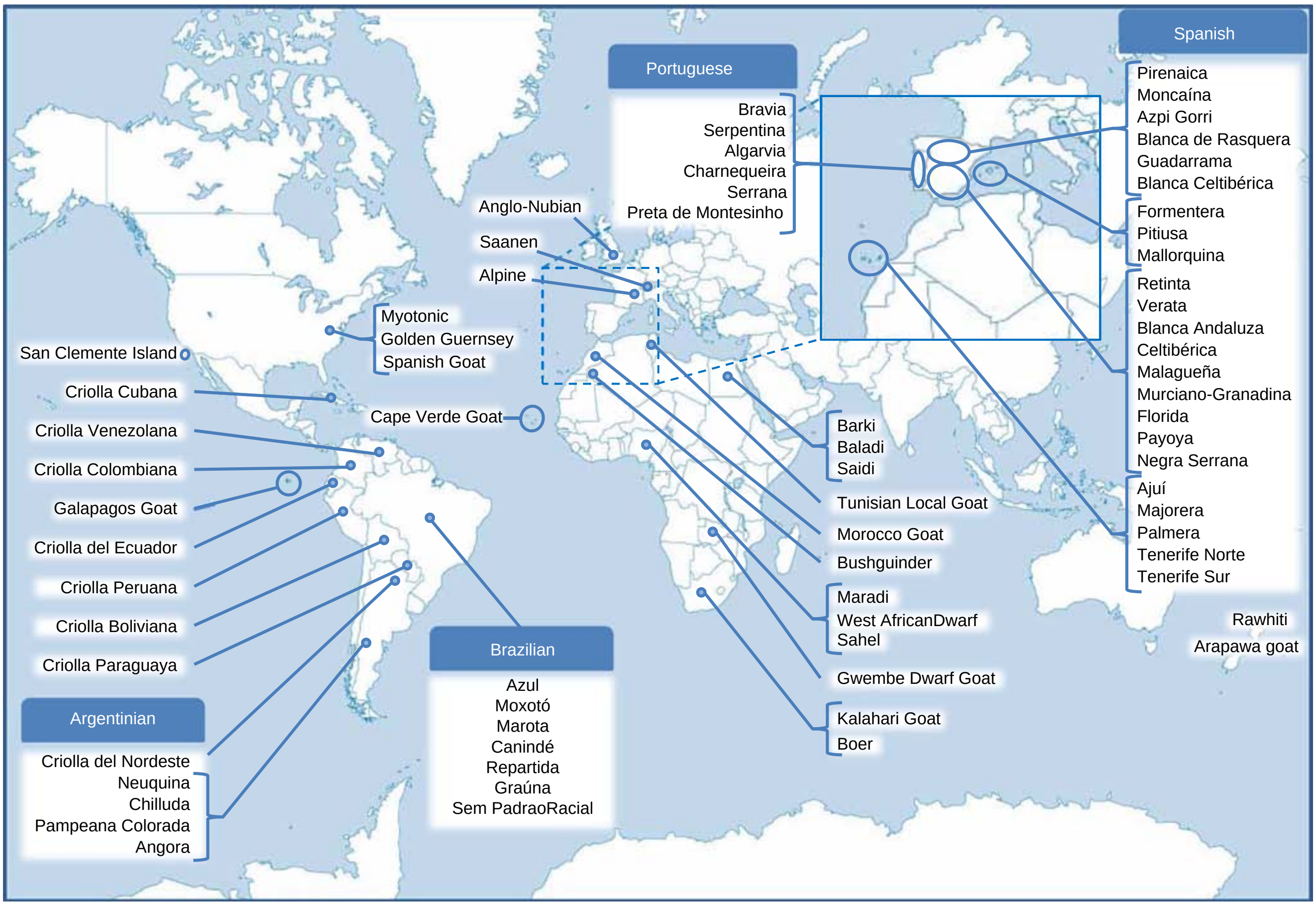

Figure 1 (colour online) World-spanning geographic distribution of the 71 caprine populations included in the study.

number of genetic clusters was assessed by both the method of Evanno *et al*. (2005) and the plots of the likelihood of the data for different *K*-values (i.e. plot of mean ln Pr(X/*K*) and its variance *v*. *K*). Genetic landscape analyses were carried out to produce synthetic contour maps built using the graphical library of the statistical software R (http://www.r-project.org) with spatial interpolation of the admixture coefficients (Q matrix) obtained with STRUCTURE for $K = 4$ on a worldwide map (http://membres-timc.imag.fr/Olivier.Francois/admix_display.html), considering the centre of geographical dispersion of each breed.

The influence of breed ancestry diversity (admixture) on the level of genetic diversity was also investigated. The degree of ancestry diversity of each breed was calculated as $1 - \Sigma(q_k)^2$, where q_k is an average fraction of the breed's genetic ancestry from the *k* separate genetic clusters at the optimal *K*, as identified with the STRUCTURE program (Tapio *et al*., 2010). The significance of admixed ancestries as sources of within-breed diversity was analysed by means of the comparison between the obtained ancestry and the within-breed diversity estimates.

Admixture proportions in Creole populations, that is the proportion of the contemporary gene pools derived from different parental clusters, were analysed using the likelihood-based method implemented in the software LEADMIX (Wang, 2003), which takes sampling variance, genetic drift and genetic differentiation between the parental populations into account. Analyses were run considering Creole breeds as a whole and by grouping populations according to the breed relationships obtained with both STRUCTURE and the Neighbour-net.

Results

Genetic diversity

In total, 71 goat populations distributed throughout the world (Figure 1) were analysed using 20 microsatellite markers, which were highly variable and allowed the detection of a mean number of about six alleles per locus per breed, with global H_e and H_o of 0.64 and 0.60, respectively, and a total of 263 alleles across all populations and loci (Supplementary Table S1). Genetic parameters and distances were computed for the main groups of breeds, defined according to their geographical location, that is, Iberian, Canarian, African and Creole breeds (Table 1). Among the four defined clusters, the African caprine populations were the most variable in terms of Ar (5.00) and H_e (0.67). Positive within-breed inbreeding coefficient (F_{IS}) values were obtained for all breed groups, with the highest values displayed by Creole (0.08) and Canarian (0.07) groups (Table 1).

The individual analysis of breeds showed that the smallest Ar ($\leqslant$3.0) was found in Golden Guernsey, Galapagos Goat,

Table 1 *Genetic diversity indicators across 20 microsatellite loci for four geographic groups of worldwide goat breeds*

Breed group	*n*	Ar ± SE	H_e ± SE	H_o ± SE	F_{IS} ± SE
Iberian	775	4.76 ± 0.10	0.653 ± 0.009	0.611 ± 0.009	0.065 ± 0.010
Canarian	200	4.52 ± 0.13	0.637 ± 0.010	0.592 ± 0.011	0.072 ± 0.015
African	441	5.00 ± 0.17	0.673 ± 0.015	0.634 ± 0.014	0.059 ± 0.009
Creole	786	4.36 ± 0.15	0.629 ± 0.016	0.578 ± 0.015	0.082 ± 0.009

n = sample size; Ar = allelic richness; H_e = expected heterozygosity; H_o = observed heterozygosity; F_{IS} = within-breed inbreeding coefficient. Full details on breed groups given in Supplementary Table S1.

Arapawa Goat and Rawhiti breeds, and the highest Ar (>5.6) in Saidi, Criolla Paraguaya and Baladi populations (Supplementary Table S1). The H_e per breed ranged between 0.37 (Golden Guernsey) and 0.77 (Criolla Paraguaya). The Spanish Verata and the Portuguese Preta de Montesinho breeds had the highest within-breed significant F_{IS} value, with estimates around 0.17 (Supplementary Table S1).

At the locus level, the number of alleles ranged between 3 (*MAF209*) and 23 (*OarFCB304*), with an average of 13.2 alleles per locus (Supplementary Table S2). The He and H_o per locus over all breeds ranged from 0.27 (*ETH225*) to 0.81 (*MM12*), and from 0.25 (*ETH225*) to 0.77 (*MM12*), respectively (Supplementary Table S2). Five microsatellite markers were in HW disequilibrium in more than six breeds ($P < 0.001$) (Supplementary Table S2).

Genetic distances and clustering

Reynolds genetic distances estimated for the various pairs of breed groups ranged from 0.027 to 0.095 for the African–Creole and Iberian–Canarian groups, respectively (Supplementary Figure S1). A strong level of genetic differentiation was observed for the Canarian breeds relative to the other groups, with a mean breed distance of 0.078 (Supplementary Figure S1). African and Creole populations displayed the lowest level of genetic divergence when compared with the other breed groups (0.044 and 0.045, respectively). The mean genetic distance among breeds of the same group was 0.16 for Creole, 0.11 for African and 0.05 for Iberian breeds (results not shown). The global mean value for F_{ST} across breeds was 0.13, which indicates that 87% of the total variability can be ascribed to within-breed variation.

The Neighbour-net dendrogram supported the existence of three major clusters, that is, Iberian, Canarian and Brazilian breeds (Figure 2), while the other breeds were not so closely grouped. Most Creole populations grouped around Iberian breeds, consistent with a close genetic relationship among all these goats, with the exceptions of the Brazilian, Colombian and Cuban breeds. The Cuban breed grouped close to the Canarian and West African populations (Bushguinder from the Saharawi Camps, Maradi, West African Dwarf and Sahel from Nigeria, Moroccan and Tunisian goats), whereas the Brazilian breeds are in a separate cluster, together with goats from Cape Verde. The Myotonic and Spanish Goat breeds from the United States, the Criolla Paraguaya and the Argentinian breeds Neuquina, Chilluda, Angora type and Pampeana Colorada, clustered at the centre of the graph, with influences from Iberia, Egypt, Zambia and South Africa. Another Creole group comprising Criolla Venezolana, del Ecuador, Peruana, Boliviana and the Argentinian del Nordeste breeds, was placed nearby Iberian populations, reflecting the influence that this ancestral group has had on their autosomal genetic pool. Isolated island (San Clement Island, Galapagos Goat, Rawhiti and Arapawa populations) and rare (Golden Guernsey) breeds showed greater genetic distances, reflected by their long branches. Regarding Iberian breeds and in accordance with previous results reported by Martínez *et al.* (2015), populations from Spain (with the exception of Canary Island breeds) and Portugal showed a weak population structure, with a strong degree of interspersing regardless of their geographical distribution.

Results from STRUCTURE are shown in Figure 3. The most likely number of genetic clusters was observed for $K = 20$ (Supplementary Figure S2). Setting *K* at 2 split the genetic origins into the Iberian and African components. $K = 3$ separated the group of Brazilian populations, which divided into two groups at higher *K*-values (Figure 3, Supplementary Table S3), indicating their high level of differentiation. The genetic structure detected for the remaining Creole breeds corroborated the major relationships highlighted by the Neighbour-net clustering, with slight differences resulting from imposing a tree-like structure on populations that are the result of complex admixture events.

The spatial interpolation of the genetic contributions to goat populations at $K = 4$ reflected geographical patterns of genetic dispersion underlying the formation of current breeds (Rosenberg *et al.*, 2002) and identified various groups (Supplementary Figure S3): Cluster I encompassing all Iberian breeds and depicting their relationship with the Creole breeds San Clemente Island, Galapagos Goat, Criolla Venezolana, Colombiana, Peruana, Boliviana and del Noroeste; Cluster II including Cape Verde Goat, Bushguinder, Nigerian and Canary Island populations, along with their Creole counterparts Criolla Cubana, Venezolana, Peruana and Boliviana; Cluster III comprising Egyptian, Tunisian, Moroccan, Zambian and South African breeds, as well as the Creole Myotonic, Spanish Goat, Criolla del Ecuador and Paraguaya, Neuquina, Chilluda, Pampeana Colorada and Angora type; and Cluster IV exclusively composed by the Brazilian breeds. In the analyses done with STRUCTURE, from $K = 6$ onwards Cluster III split into its North and South African components, and only the main relationship between the North African breeds (Barki, Baladi, Saidi,

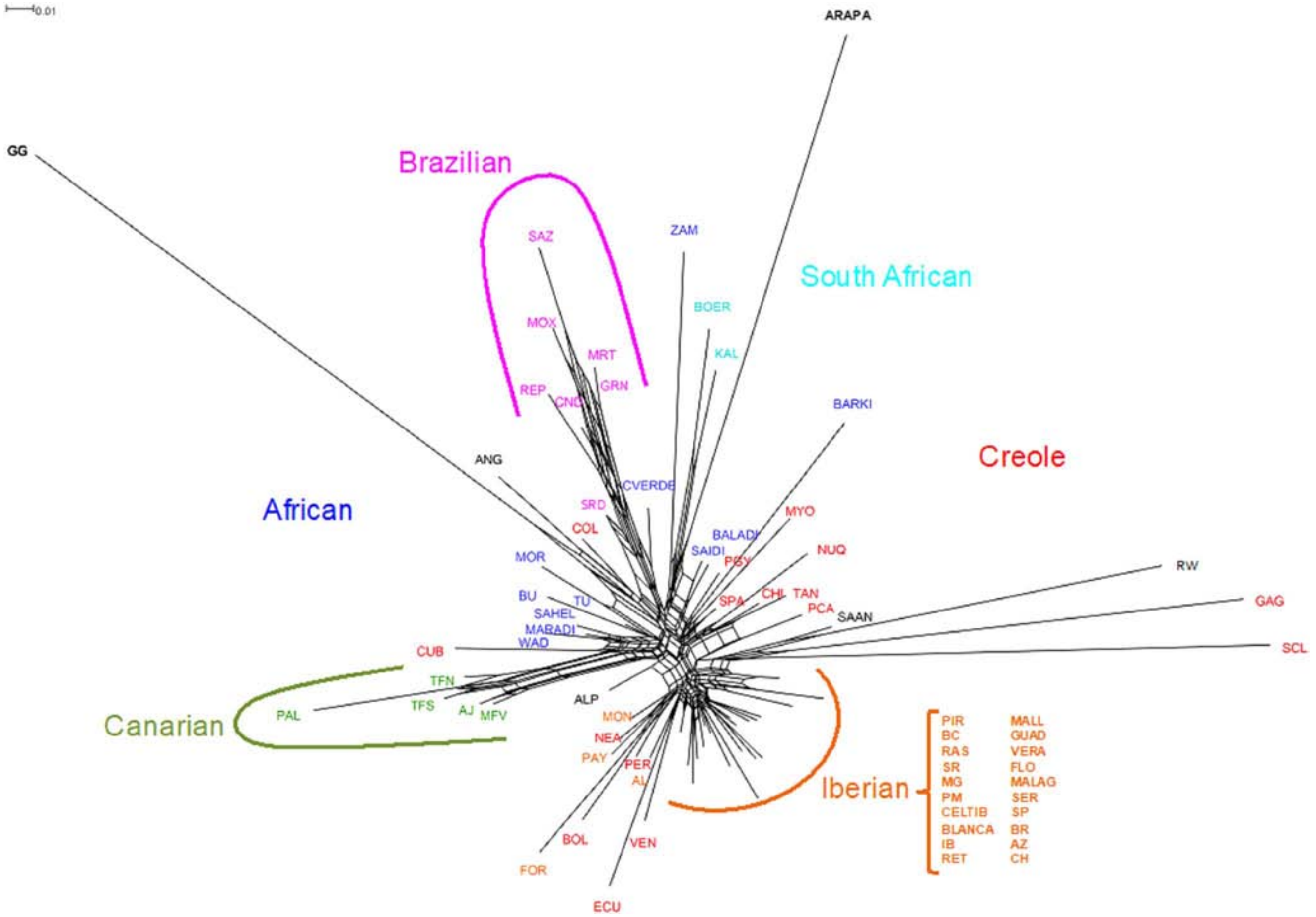

Figure 2 (colour online) Neighbour-net dendrogram constructed from the Reynolds genetic distances among 71 goat populations. For more information on breeds, see Supplementary Table S1.

Morocco Goat and Tunisian Local Goat) and the Creole populations remained.

Although the representation of STRUCTURE outputs helps to visualise the different genetic origins implicated in the formation of Creole breeds, it must be kept in mind that these results are highly sensitive to the model assumptions and the breeds included in the analysis, which is reflected in the erratic cluster assignment shown by some populations for different runs at the same or different *K*-values (Figure 3), for example, the case of the Galapagos Goat.

When the most likely number of inferred genetic origins is assumed ($K = 20$), only 26 out of the 71 analysed breeds had a contribution from one of the clusters above 0.7 (Supplementary Table S3), which suggests the existence of important levels of admixture or an incomplete divergence of populations. Breed H_e and the level of admixture based on STRUCTURE results for $K = 20$ (Supplementary Table S3) were positively correlated ($r = 0.7$, $P < 0.0001$), indicating that the level of admixture of a breed may explain a high proportion of its estimated genetic variability (i.e. around 50% of the expected genetic variability can be ascribed to the level of ancestry diversity resulting from admixture).

FCA clustering methods were used to further analyse the genetic structure of the 71 caprine breeds. The first two axes explained 15.3% of the variability, contributing with 8.3% and 7% of the total inertia, respectively (Figure 4a). The first axis clearly separated Brazilian populations from the remaining breeds, such that the Brazilian group accounted by itself for 71.9% of the total inertia in this axis. In axis 2, all breeds accounted for <10% of inertia, with the two South African populations being the most differentiated (16.2%). When Brazilian goats were excluded in a second analysis (Figure 4b), axes 1 and 2 contributed with a similar percentage of the total inertia as before (8.3% and 7.2%, respectively), but the relationships among the remaining breeds could be better assessed. In this case, all Iberian breeds grouped in a tight cluster whereas Creole and African populations appeared mixed, with the major exceptions of South African breeds, the Barki breed from Egypt and the Canary Island populations. As previously shown, the Criolla Cubana had a closer proximity to the Canary Islands, Cape Verde and Nigeria populations, and the position of the remaining Creole breeds reflected different influences from African and Iberian groups.

Ancestral genetic contributions to Creole goats

LEADMIX was used for estimating the genetic contributions to and degree of admixture of the Creole populations. The model underlying this method assumes that a distant ancestral population split into the different parental

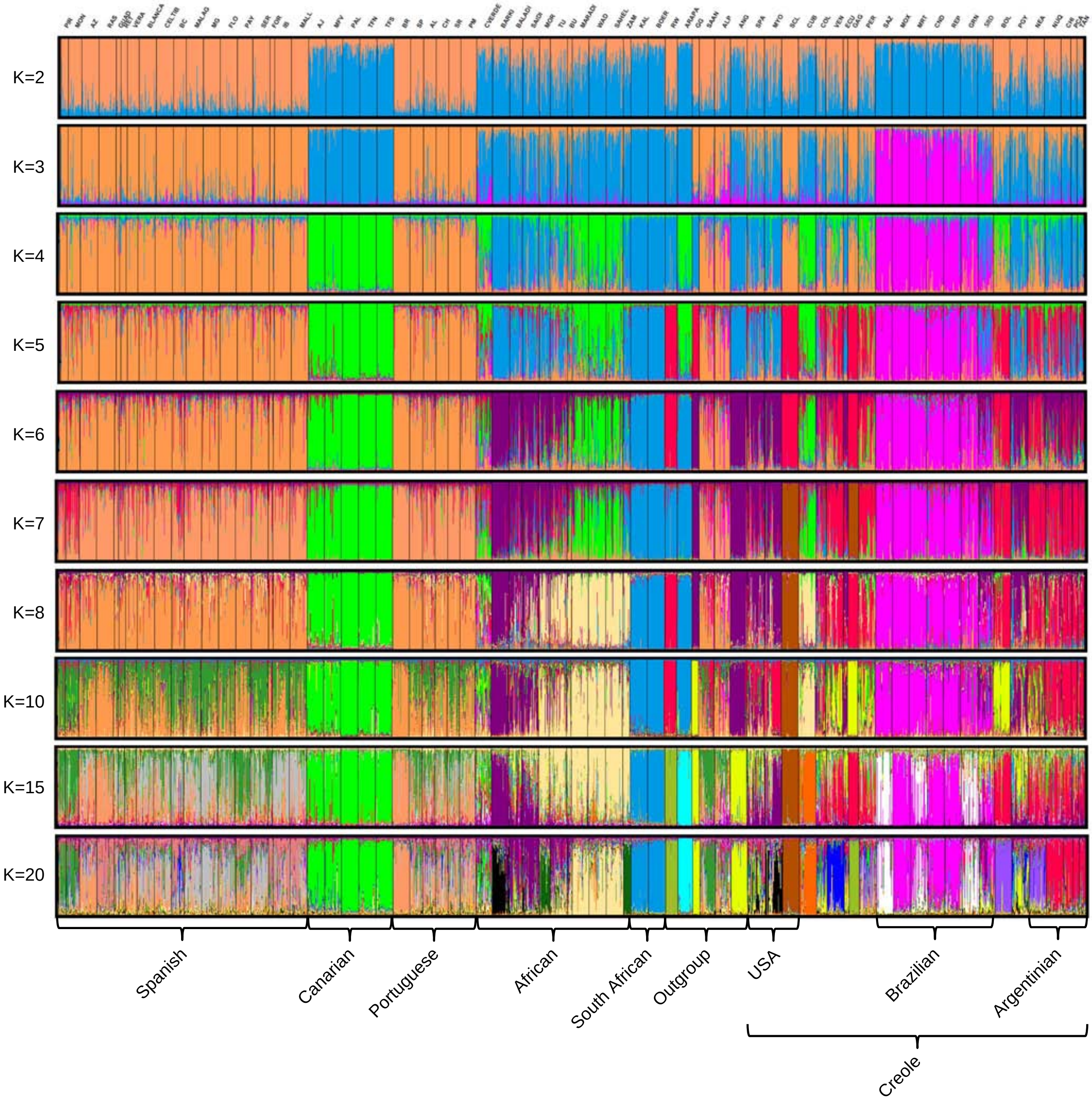

Figure 3 (colour online) Estimated membership coefficients (q) of goat populations assuming *K*-values from 2 to 20 as inferred with STRUCTURE. For more information on breeds, see Supplementary Table S1.

populations, which diverge for some period of time and then converge once to form the admixed population. The goal in our analyses was to estimate the genetic contribution from various parental populations to the Creole gene pool, which was considered either as a whole group or as various subgroups. Based on the results from the spatial interpolation of the Q matrix obtained with STRUCTURE for $K = 4$ (Supplementary Figure S3) and the Neighbour-net (Figure 2), three potential ancestral population groups were defined (Figure 5): (i) Cluster I, including the Iberian breeds; (ii) Cluster II, encompassing the breeds from Cape Verde, Saharawi Camps, Nigeria and Canary Islands; and (iii) Cluster III, which gathers the Egyptian, Tunisian, Moroccan, Zambian and South African breeds. The admixture estimates indicated that the whole group of Creole goat populations, including Brazilian breeds, had average ancestral contributions of 40%, 26% and 34% from Clusters I, II and III, respectively (Figure 5).

The STRUCTURE and Neighbour-net analyses supported the existence of various Creole groups, which may reflect variable proportional contributions from the different ancestral genetic groups. Therefore, estimations with LEADMIX for each of the three main Creole groups, as well as for

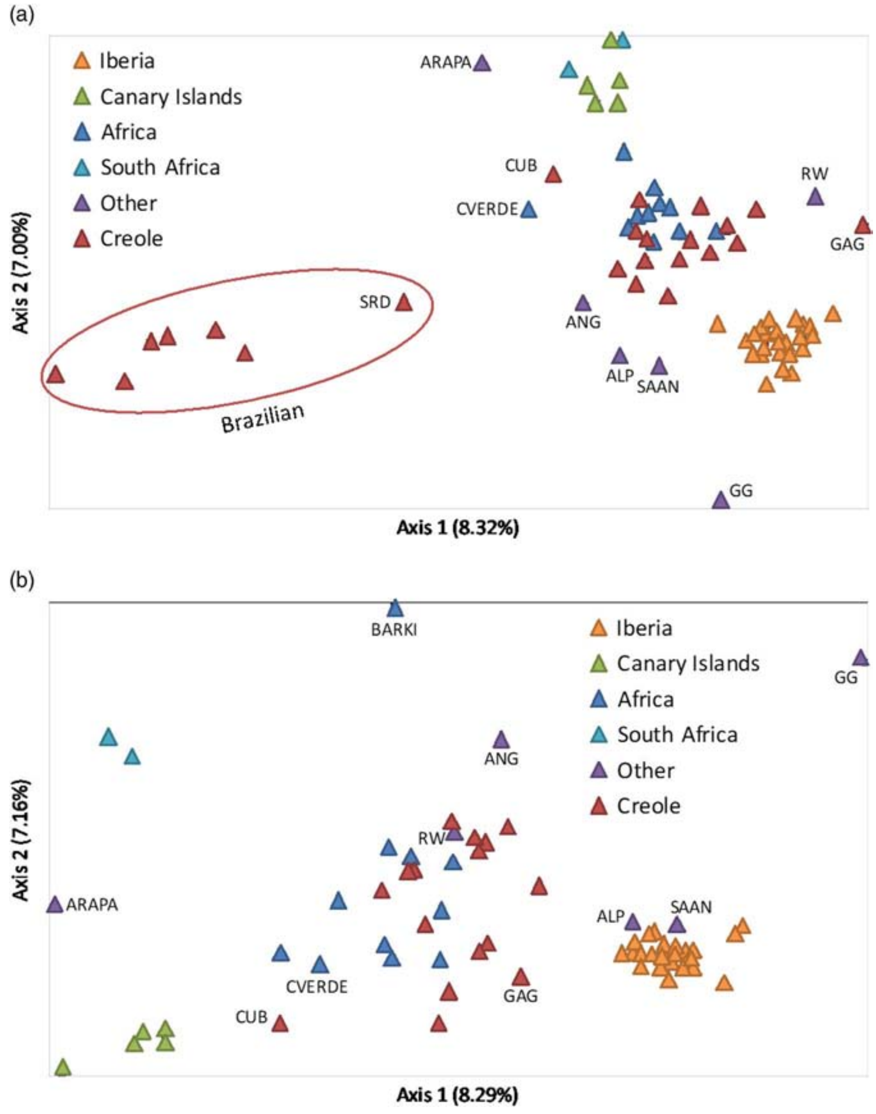

Figure 4 (colour online) Graphical representation of the tridimensional factorial correspondence analysis. (a) 71 goat populations, (b) 64 goat breeds after removing the Brazilian populations. For more information on the groups of breeds, see Supplementary Data S1.

the Criolla Cubana, were also performed (Figure 5). Brazilian breeds were excluded from this analysis given their lack of consistent relationships with any of the ancestral origins analysed. Although our admixture estimates could be distorted by admixture from other genetic origins that may have not been included in this study, generally speaking the results indicated that 55% of the autosomal gene pool of Group 1 (San Clemente Island, Criolla Colombiana, Galapagos goat and Criolla del Nordeste populations) was of Iberian origin and about 45% was contributed by Cluster II. In Group 2 (Spanish Goat, Myotonic, Criolla del Ecuador, Criolla Paraguaya, Neuquina, Chilluda, Pampeana Colorada and Angora type), contributions from all parental origins could be identified, with nearly the same influence from Cluster I and III (40%) and a lower contribution from Cluster II (20%). Group 3 (Criolla Venezolana, Peruana and Boliviana) also displayed a major Iberian influence (56%), with the other major contribution to their genetic pool from Cluster II (44%). Criolla Cubana showed a clear predominance of Cluster II genetic contribution (78%) and a minor influence from Iberian breeds (22%).

Discussion

The low phylogeographic structure of goat populations observed in our study is a recurring theme when analysing genetic variation of caprine populations, where genetic variation at the within-population level usually ranges between 85% and 93% (Li *et al.*, 2002; Cañón *et al.*, 2006; Ribeiro *et al.*, 2012; Wei *et al.*, 2014). In agreement with this general trend, our results showed that 87% of the genetic variation corresponds to the within-population component, meaning that the breeds included in this study are poorly differentiated at the genetic level. However, the mean genetic distances among goat breeds of the Creole group (0.16) are higher than those obtained among breeds of the Iberian (0.05)

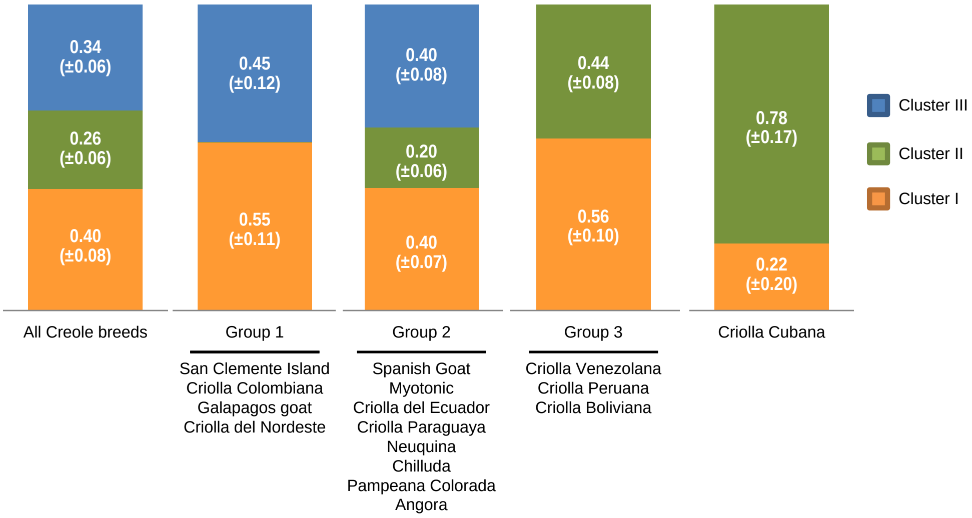

Figure 5 (colour online) Ancestral genetic contributions from clusters I, II and III to Creole goats. Creole breeds were considered as a whole as well as grouped in four different clusters. SD (between brackets) obtained from 1000 bootstraps over loci. For more information on the ancestral populations, see Supplementary Data S1.

and African (0.11) groups, which may reflect a different history in American goats, with genetic contributions from several parental populations and genetic drift, as well as admixture with other breeds over time.

In general, Creole populations lack a clear geographical pattern of genetic diversity, with populations of different admixed ancestry sharing relatively close locations all over the large geographical range of the Creole samples included in this study. Thus, we find breeds of mainly Iberian and Cluster III African origin (San Clemente Island, Criolla Colombiana, Galapagos Goat and Criolla del Nordeste) interspersed with populations of Iberian and Cluster II African origin (Criolla Venezolana, Peruana and Boliviana), or with breeds displaying signatures of admixture from these three ancestral origins (Spanish Goat, Myotonic, Criolla del Ecuador and Paraguaya, Neuquina, Chilluda, Pampeana Colorada and Angora type). This contrasts with the results obtained for European, African and Asian goat breeds (Cañón *et al.*, 2006; Pereira *et al.*, 2009; Wei *et al.*, 2014). However, the lack of a clear geographical pattern in American populations could be anticipated, given the different model of breed expansion that took place in this continent. Whereas in the Old World the dissemination of livestock species has, with some exceptions, been slow, continuous and mostly by geographical proximity (Pereira *et al.*, 2009), the introduction of different domesticated species in America is believed to have followed a rapid and discontinuous progression, with multiple waves of restocking introductions from Iberia and the Atlantic archipelagos that have disseminated and admixed breeds randomly (Rodero *et al.*, 1992; Amills *et al.*, 2009).

The lack of a remarkable genetic signature from the occidental Old World in Brazilian goats is in strong contrast with the distinctive Iberian and African footprints found in the remaining American goat populations. Nevertheless, some influence of Cape Verde goats was detected in Brazilian breeds, which would reflect the importance of this archipelago as a layover port in the journey of Portuguese between the Iberian Peninsula and the Americas (Primo, 2004).

Opposed to the minor genetic signatures of admixed ancestry between African and Creole cattle (Gautier and Naves, 2011; Martínez *et al.*, 2012), the large contribution of the African gene pool to nearly all Creole goat populations revealed in our study (Figures 2 to 5) does not seem to be entirely explained by the existence of migration between Northwest Africa and Iberia, as the impact of African introgression into Iberian breeds seems to be quite limited (Pereira *et al.*, 2009; Martínez *et al.*, 2016). Even though a scenario of multiple introductions of goats from the Atlantic archipelagos may be plausible based on historical and genetic evidences, it does not explain the ancestry component from breeds originating in other African regions detected in our study. Alternatively, livestock could have been brought directly from Africa in slave trade ships or via the ships of the 'Carreira da India', which allowed a continuous trade flow from India to the East Coast of Africa and then to Brazil (O'Flanagan, 2008). The important contribution of the African gene pool is particularly patent in the Criolla Cubana breed, and if we consider it as a remnant of the first goats brought from Spain, a predominant West African influence is still visible, which might have been exported to the rest of the Americas by these populations.

Regarding the Iberian ancestry of Creole goats, we find an important autosomal Iberian component in nearly all Creole populations studied, with the exception of Brazilian breeds. This result is in contrast with the one reported by Amills *et al.* (2009), who assessed mitochondrial DNA in various goat populations and failed to find a clear Iberian signature in some Creole populations, probably because this marker reflects only the history of female-mediated gene flow and does not predict nuclear genetic diversity. As previously reported by Martínez *et al.* (2015), the gene pool of Spanish and Portuguese goats cannot be separated nowadays, and even less in the 16th century, when the two countries were under the same rule and selective breeding and breed differentiation were absent from caprine herds. This is reflected in the mainly equidistant relationships of Creole breeds relative to Iberian populations, regardless of their Spanish or Portuguese origin.

As expected, there was a close genetic relationship between the Canary Island breeds and West African goats, supporting the idea that Canary goats have suffered an important founder effect nearly 2200 years ago when the first goat flock of African roots was brought to the archipelago by Berber populations (Rodero *et al*., 1992; Ferrando *et al*., 2015). This would explain the high differentiation of Canary populations relative to the continental Iberian goat breeds. We have also found some degree of genetic affinity between caprine populations from Northwest Africa (Morocco and Tunisian Local goats) and the Middle East goats represented by Egyptian breeds, in agreement with Pereira *et al*. (2009) and Martínez *et al*. (2016), whereas Bushguinder from the Saharawi Camps seems more related to West African populations. The Cape Verdean goats displayed a close proximity to Canary Island breeds, both in STRUCTURE and FCA analysis, which is consistent with previous genetic studies (Pereira *et al*., 2009) and historical records that report the introduction of goats from the Canary Islands, Africa and Portugal in Cape Verde in the 15th century (Rodero *et al*., 1992).

In conclusion, the genetic relationships between Creole goats and their main presumed genetic origins were investigated with data on neutral autosomal genetic markers from >2000 samples representing a wide range of Iberian, African and Creole breeds. Most of the genetic diversity observed in Creole goat populations derives from admixture between Iberian and African lineages, making them an important reservoir to be managed and improved. The main insights that summarise this work are: (i) the lack of geographical patterns in the American distribution of goat genetic diversity; (ii) the wide differentiation among various Creole breeds; (iii) the confirmation of the important autosomal Iberian signatures in nearly all Creole populations studied; (iv) the surprisingly high contribution of the African gene pool to Creole breeds, especially to Criolla Cubana; and finally (v) the particular genetic isolation of Brazilian breeds, with some influence from Cape Verde goats. These results provide a comprehensive characterisation of goat genetic diversity in the Americas, and a dissection of the Iberian and African components that gave origin to different Creole caprine breeds, which will be useful for the recognition of these breeds and the development of appropriate selection and conservation programmes aimed at their sustainable use, thus preventing the loss of genetic diversity.

Acknowledgements

The authors thank 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 has been partially supported by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) RZ2004-00037-00-00, RZ2006-00005-C02-00, RZ2007-00005-C02-01 and RZ2007-00005-C02-02 projects. C.G. was supported by a contract grant from the Fundação para a Ciência e a Tecnologia (IF/00866/2014), Portugal. C.B.S. was supported by a grant from the Fundação para a Ciência e a Tecnologia (Ref. SFRH/BM/1770/2004).

Declaration of interest

The authors declare that they have no competing interests

Ethics statement

Not applicable.

Software and data repository resources

Not applicable.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1751731117003627

References

Amills M, Ramírez O, Tomàs A, Badaoui B, Marmi J, Acosta J, Sànchez A and Capote J 2009. Mitochondrial DNA diversity and origins of South and Central American goats. Animal Genetics 40, 315–322.

Belkhir K, Borsa P, Chikhi L, Raufaste N and Bonhomme F 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Université de Montpellier II. Retrieved on 15 November 2015 from http://kimura.univ-montp2.fr/genetix/

Cañón J, García D, García-Atance MA, Obexer-Ruff G, Lenstra JA, Ajmone-Marsan P, Dunner S and ECONOGENE Consortium 2006. Geographical partitioning of goat diversity in Europe and the Middle East. Animal Genetics 37, 327–334.

Evanno G, Regnaut S and Goudet J 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611–2620.

FAO Commission on Genetic Resources for Food and Agriculture 2011. Molecular genetic characterization of animal genetic resources. FAO Animal Production and Health Guidelines. No 9, Rome. Retrieved 10 December 2015 from http://www.fao.org/docrep/014/i2413e/i2413e00.pdf.

Felsenstein J 2004. Inferring phylogenies. Sinauer Associates Inc., Sunderland, MA, USA.

Ferrando A, Manunza A, Jordana J, Capote J, Pons A, Pais J, Delgado T, Atoche P, Cabrera B, Martínez A, Landi V, Delgado JV, Argüello A, Vidal O, Lalueza-Fox C, Ramírez O and Amills M 2015. A mitochondrial analysis reveals distinct founder effect signatures in Canarian and Balearic goats. Animal Genetics 46, 452–456.

Gautier M and Naves M 2011. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. Molecular Ecology 20, 3128–3143.

Ginja C, Gama LT, Martínez A, Sevane N, Martin-Burriel I, Lanari MR, Revidatti MA, Aranguren-Méndez JA, Bedotti DO, Ribeiro MN, Sponenberg P, Aguirre EL, Alvarez-Franco LA, Menezes MPC, Chacón E, Galarza A, Gómez-Urviola N, Martínez-López OR, Pimenta-Filho EC, da Rocha LL, Stemmer A, Landi V and Delgado-Bermejo JV 2017. Genetic diversity of Creole goats and patterns of population structure reflect the mode of goat colonization in the Americas. Animal Genetics 48, 315–329.

Ginja C, Cortés O, Gama LT, Delgado JV, Amills M, Bruno de Sousa C, Cañón J, Capote J, Dunner S, Ferrando A, Gómez Carpio M, Jordana J, Landi V, Manunza A, Martin-Burriel I, Pons Barro A, Rodellar C, Santos-Silva F, Sevane N, Vidal O, Zaragoza P, Martínez AM (2018). Conservation of goat populations from Southwestern Europe based on molecular diversity criteria. In Sustainable goat production in adverse environments – welfare, health and breeding. (ed. J Simões and C Gutierrez), Vol. I. Chapter 29. Springer (in press).

Goudet J 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Lausanne University. Retrieved 4 March 2015 from http://www2.unil.ch/popgen/softwares/fstat.htm

Huson DH and Bryant D 2006. Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23, 254–267.

Li MH, Zhao SH, Bian C, Wang HS, Wei H, Liu B, Yu M, Fan B, Chen SL, Zhu MJ, Li SJ, Xiong TA and Li K 2002. Genetic relationships among twelve Chinese indigenous goat populations based on microsatellite analysis. Genetics Selection and Evolution 34, 729–744.

Martínez AM, Gama LT, Cañón J, Ginja C, Delgado JV, Dunner S, Landi V, Martín-Burriel I, Penedo MC, Rodellar C, Vega-Pla JL, Acosta A, Alvarez LA, Camacho E, Cortés O, Marques JR, Martínez R, Martínez RD, Melucci L, Martínez-Velázquez G, Muñoz JE, Postiglioni A, Quiroz J, Sponenberg P, Uffo O, Villalobos A, Zambrano D and Zaragoza P 2012. Genetic footprints of Iberian cattle in America 500 years after the arrival of Columbus. PloS One 7, e49066.

Martínez AM, Gama LT, Delgado JV, Cañón J, Amills M, de Sousa CB, Ginja C, Zaragoza P, Manunza A, Landi V, Sevane N and BioGoat Consortium 2015. The Southwestern fringe of Europe as an important reservoir of caprine biodiversity. Genetics Selection and Evolution 47, 86.

Martínez A, Manunza A, Delgado JV, Landi V, Adebambo A, Ismaila M, Capote J, El Ouni M, Elbeltagy A, Abushady AM, Galal S, Ferrando A, Gómez M, Pons A, Badaoui B, Jordana J, Vidal O and Amills M 2016. Detecting the existence of gene flow between Spanish and North African goats through a coalescent approach. Scientific Reports 6, 38935.

Naderi S, Rezaei HR, Taberlet P, Zundel S, Rafat SA, Naghash HR, el-Barody MA, Ertugrul O, Pompanon F and Econogene Consortium 2007. Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. PloS ONE 2, e1012.

O'Flanagan P 2008. Port cities of Atlantic Iberia, c. 1500–1900. Ashgate Publishing Limited, Surrey, UK.

Park SDE 2001. Trypanotolerance in west African cattle and the population genetic effects of selection. University of Dublin, Ireland.

Pereira F, Queirós S, Gusmão L, Nijman IJ, Cuppen E, Lenstra JA, Econogene Consortium, Davis SJ, Nejmeddine F and Amorim A 2009. Tracing the history of goat pastoralism: new clues from mitochondrial and Y chromosome DNA in North Africa. Molecular Biology and Evolution 26, 2765–2773.

Primo AT 2004. América: conquista e colonização. Movimento, Porto Alegre, Brazil.

Pritchard JK, Stephens M and Donnelly P 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Raymond M and Rousset F 1995. GENEPOP (Version 1.2): population genetics software for exact test and ecumenicism. Journal of Heredity 86, 248–249.

Ribeiro MN, Bruno-de-Sousa C, Martinez-Martinez A, Ginja C, Menezes MP, Pimenta-Filho EC, Delgado JV and Gama LT 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 JV and Rodero E 1992. Primitive Andalusian livestock and their implications in the discovery of America. Archivos de Zootecnia 41, 383–400.

Rosenberg NA 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4, 137–138.

Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA and Feldman MW 2002. Genetic structure of human populations. Science 298, 2381–2385.

Tapio M, Ozerov M, Tapio I, Toro MA, Marzanov N, Cinkulov M, Goncharenko G, Kiselyova T, Murawski M and Kantanen J 2010. Microsatellite-based genetic diversity and population structure of domestic sheep in northern Eurasia. BMC Genetics 10, 11–76.

Walsh PS, Metzger DA and Higushi R 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10, 506–513.

Wang J 2003. Maximum-likelihood estimation of admixture proportions from genetic data. Genetics 164, 747–765.

Wei C, Lu J, Xu L, Liu G, Wang Z, Zhao F, Zhang L, Han X, Du L and Liu C 2014. Genetic structure of Chinese indigenous goats and the special geographical structure in the Southwest China as a geographic barrier driving the fragmentation of a large population. PLoS One 9, e94435.

Weir BS and Cockerham CC 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358.