Livestock Science 197 (2017) 1–7

Contents lists available at ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

# Contributions to diversity rather than basic measures of genetic diversity characterise the spreading of donkey throughout the American continent

J. Jordana[a,*], F. Goyache[b], A. Ferrando[a], I. Fernández[b], J. Miró[c], A. Loarca[d], O.R. Martínez López[e], J.L. Canelón[f], A. Stemmer[g], L. Aguirre[h], M.A.C. Lara[i], L.A. Álvarez[j], S. Llambí[k], N. Gómez[l], L.T. Gama[m], R.D. Martínez[n], E. Pérez[o], A. Sierra[p], M.A. Contreras[q], V. Landi[r], A. Martínez[r], J.V. Delgado[r]

[a] Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[b] SERIDA-Deva, Área de Genética y Reproducción Animal, Camino de Rioseco 1225, 33394 Gijón, Spain
[c] Departament de Medicina i Cirurgia Animals, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[d] Ministerio de Agricultura, Ganadería y Alimentación, Quetzaltenango, Guatemala
[e] Centro Multidisciplinario de Investigaciones Tecnológicas, Dirección General de Investigación Científica y Tecnológica, Universidad Nacional de Asunción, Asunción, Paraguay
[f] Departamento de Producción e Industria Animal, Decanato de Ciencias Veterinarias, Universidad Centroccidental Lisandro Alvarado, Barquisimeto, Venezuela
[g] Facultad de Ciencias Agrícolas y Pecuarias, Universidad Mayor de San Simón, Cochabamba, Bolivia
[h] Centro Biotecnología Reproductiva Animal, Universidad Nacional de Loja, Loja, Ecuador
[i] Instituto de Zootecnia, 13.460-000 Nova Odessa, SP, Brazil
[j] Universidad Nacional de Colombia, Sede Palmira, Cra. 32 No 12-00, Palmira, Valle del Cauca, Colombia
[k] Instituto de Producción Animal, Área Genética, Facultad de Veterinaria, Universidad de la República (UdelaR), CP11600 Montevideo, Uruguay
[l] Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Micaela Bastidas de Apurímac, Abancay, Perú
[m] CIISA - Faculdade de Medicina Veterinária, Universidade de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal
[n] Mejora y Conservación de Recursos Genéticos, Facultad de Ciencias Agrarias, Universidad Nacional de Lomas de Zamora, Ruta 4 - Km. 2., 1836 Llavallol, Argentina
[o] Departamento de Salud y Explotación Animal, Universidad de Granma, 85100 Bayamo, Cuba
[p] Posgrado e Investigación, Instituto Tecnológico de Conkal, Yucatán, México
[q] Escuela de Medicina Veterinaria, UST Sede Santiago, Santiago, Chile
[r] Departamento de Genética, Facultad de Veterinaria, Universidad de Córdoba, 14071 Córdoba, Spain

## ARTICLE INFO

*Keywords:*
Contributions to diversity
Geographic patterns
Gene diversity
Allelic richness
Microsatellites

## ABSTRACT

Donkey was introduced into the Americas soon after its discovery in the 15th century. However, there is no historical consensus on how they spread across the continent. In a previous study, two distinct genetic pools (Clusters A -Southern part - and B - Northern part of South America and Central America) were identified, with likely confluence in Colombia. The aim of this study was to evaluate whether the main genetic diversity parameters, such as gene diversity (GD) and allelic richness (k), or the relative contributions of various breeds to these parameters are useful indicators to give genetic support to historical information on putative routes of the spreading of donkeys across the American continent. In full agreement with historical sources suggesting that Greater Antilles were the first breeding nucleus, both total contributions to gene diversity ($gGD_T$) and to allelic richness ($C_T^{(k)}$) showed a higher ability to identify the "abundant centre" of the species on the Continent. Even though there are historical reports suggesting various entry points of the donkey into the continent (e.g. in Brazil), these parameters suggested that, in our dataset, the Cuban donkey population was the more likely representative of the first breeding nucleus of the species. Central and South American donkey populations in the surroundings of the Caribbean Gulf would more likely be early derivatives of Antillean donkey. The strong North-South genetic structure was confirmed for the American donkey metapopulation. Current analyses suggest that populations classified into Cluster A (South) are essentially a sample of the genetic background of Cluster B (North). The Andean route had the highest importance in the formation of the South American populations. The extinction of either population belonging to Cluster B could lead to a decrease in overall genetic diversity both at the gene diversity level (negative $gGD_T$ values) and the allelic richness level (positive $C_T^{(k)}$ contributions). The opposite pattern is found for populations belonging to Cluster A. The extinction of the

* Corresponding author.
*E-mail address:* Jordi.Jordana@uab.cat (J. Jordana).

http://dx.doi.org/10.1016/j.livsci.2016.12.014
Received 8 June 2016; Received in revised form 20 December 2016; Accepted 30 December 2016

populations belonging to Cluster B would decrease the overall American donkey gene diversity in roughly 8% and would dramatically affect the number of alleles in the metapopulation (19.1%). However, the extinction of the donkey populations classified into Cluster A would increase overall gene diversity by 2.2%. Although, the genetic scenario of each individual population varies substantially, the joint conservation of the donkey populations classified into both Clusters A and B is highly advised.

## 1. Introduction

The introduction of domestic donkeys into America is well documented. After the initial establishment of founding populations in Greater Antilles the species was introduced into the mainland Continent through Mexico, to be spread northwards, and Panama to connect with commercial routes to Colombia, Venezuela and northern Brazil and southerly to Northern Andean countries such as Ecuador and Peru. From here, the species is expected to have been involved in the active trade routes between the Peruvian plateau and Pampas region (Brookshier, 1974; Laguna, 1991; Santos et al., 1992; Sponenberg, 1992; Rodero et al., 1992; Yanes, 2005; Delgado et al., 2010).

Recently, Jordana et al. (2016) analysed the American donkey population, together with Iberian and other Mediterranean donkey breeds, to ascertain their genetic structure and to identify the most likely ancestral donor populations. The authors reported the presence of two distinct genetic clusters (named A and B) in American donkey. Cluster A, formed by Southernmost American donkey populations, showed a very low genetic diversity probably subsequent to an older founder event and no significant influence of recent gene flow from Europe. Cluster B, mainly formed of donkey populations surrounding the Caribbean Gulf, showed higher polymorphism though it was not possible to reject the existence of modern gene flow from Iberian donkeys.

The ascertainment of those two genetic clusters was consistent with the historical information suggesting that the species moved from the initial Greater Antilles stock into the geographical areas surrounding the Caribbean Sea. Later on, a breeding nucleus was created in the Peruvian Plateau, from which Southernmost American donkey populations could have been formed (Laguna, 1991; Yanes, 2005). However, the indication of multiple different local genetic events due to different recent histories in the analysed populations prevented the ascertainment of the most likely routes for the spreading of the species throughout the continent (Jordana et al., 2016).

In the wild, in scenarios where gene sources and expansion patterns are known, gene diversity (expected heterozygosity; $H_e$) and allelic richness adjusted for sample size (rarefacted number of alleles per locus; $k_{(n)}$) have shown to be superior to other diversity measures to deal with the task of defining conservation priorities (Comps et al., 2001; Petit et al., 2003). Actually, those parameters have different sensitivity to stochastic processes (Eckert et al., 2008) and are able to differentiate between geographical areas acting as genetic sources (abundant centres), colonised zones and, furthermore, contact zones in which some variability parameters can show “artificially” increased values (Comps et al., 2001). From a practical point of view, both gene diversity and number of alleles per locus have the advantage of straightforward interpretation, because gene diversity illustrates the existence of balanced allelic frequencies in a population and allelic richness can characterise the degree of genetic uniqueness or distinctiveness of a population (Petit et al., 1998; Caballero and Toro, 2002). In a metapopulation both parameters can be decomposed into within- and between-population fractions (Petit et al., 1998; Caballero and Toro, 2002), therefore allowing accounting for recent local events affecting genetic signals in the populations studied. The aim of this research was to assess the usefulness of two methods for the estimation of genetic contributions to diversity, to ascertain if historical information on the routes of spreading of donkey across the American continent has genetic support. Furthermore, the potential genetic consequences of losing the populations assessed will be discussed.

## 2. Material and methods

### 2.1. Data available

The 350 American donkey DNA samples, obtained in 13 different countries (Mexico, Guatemala, Cuba, Venezuela, Colombia, Ecuador, Peru, Bolivia, Paraguay, Chile, Argentina, Uruguay and Brazil; Table 1), analysed in Jordana et al. (2016) using 14 microsatellites (VHL20, AHT4, HMS7, AHT5, HMS6, HTG10, HTG7, HMS2, HTG4, HTG6, HMS3, HTG15, HMS5 and ASB23) were available. The geographical location of sampled animals is shown in Fig. S1. Sampling strategies are described in detail in Jordana et al. (2016). DNA extraction and genotype scoring were carried out following Jordana et al. (2016) and Aranguren-Méndez et al. (2001). Following Jordana et al. (2016), for descriptive purposes samples were grouped when necessary into Cluster A (formed of Ecuador, Peru, Bolivia, Paraguay, Chile, Argentina and Uruguay samples) and Cluster B (formed of Mexico, Guatemala, Cuba, Venezuela, Colombia and Brazil samples).

### 2.2. Genetic diversity analyses

Statistical analyses were carried out using the software MolKin (current version v3.1; Gutiérrez et al., 2005). Parameters characterising genetic diversity, such as expected heterozygosity ($H_e$; Nei, 1987), heterozygote deficiency due to population inbreeding or subdivision ($F_{IS}$; Nei, 1987), raw and rarefacted ($k_{(n)}$; Hurlbert, 1971) average number of alleles per locus, and the between-population Nei's minimum distance (Nei, 1987) and molecular coancestry (Caballero and Toro, 2002) were computed. To avoid bias due to low and unequal sample sizes, parameters listed above, except for rarefacted allelic richness, were adjusted for sampling size following Cervantes et al. (2011) using as sample size the harmonic mean of the national donkey populations available (23). For the same purposes, statistical significance of the computed parameters was assessed by bootstrapping using 1000 samples and fitting sample size to 23 individuals per population. In turn, the rarefacted average number of alleles per locus was adjusted to 24 copies ($k_{(24)}$), which is twice the minimum number of individuals within a population with genotype known for all the microsatellites to allow a direct between-populations comparison of the results presented. See the MolKin User's Guide (freely available at http://www.ucm.es/info/prodanim/html/JP_Web.htm) for a detailed description of the methodologies used).

Using also MolKin, contributions to diversity were assessed following Caballero and Toro (2002) and Petit et al. (1998). Caballero and Toro (2002) proposed setting priorities for conservation using the maintenance of the maximum overall Nei's (1987) gene diversity (GD) in the preserved set of breeds as the criterion. Notice that this is equivalent to minimising the overall molecular coancestry ($\overline{f}$) because $GD = 1 - \overline{f}$. Therefore, the average GD of a given population depends on the within-subpopulation coancestry and its average distance relative to other subpopulations. This allowing the contributions to the total GD to be separated due to the within-breed diversity ($f_{ii}$) and the between-breed genetic distance. In this scenario, $GD_T = GD_W + GD_B$, where $GD_T$ is the total contribution to GD, $GD_W$ is the contribution to the within-breeds diversity and $GD_B$ the contribution to the

**Table 1**
Number of samples (N), within-population expected heterozygosity ($H_e$), deficit of heterozygotes due to populations subdivision ($F_{IS}$) rarefacted (to 24 copies) and contributions (in percentage) to within-population, between-population and total diversity of each of the thirteen analysed American donkey populations, assessed via Nei's gene diversity and rarefacted allelic richness. Abbreviations mean the following: $gGD_W$: contribution to within-population gene diversity; $gGD_B$: contribution to between-population gene diversity; $gGD_T$: total contribution to gene diversity; $C_W^{24}$: contribution to within-population allelic richness; $C_B^{24}$: contribution to between-population allelic richness; $C_T^{24}$: total contribution to allelic richness.

| Populations | N | $H_e$[a] | $F_{IS}$[a] | $k_{(24)}$ | $gGD_W$ | $gGD_B$[b] | $gGD_T$ | $C_W^{24}$ | $C_B^{24}$ | $C_T^{24}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Cluster B | | | | | | | | | | |
| Mexico | 14 | 0.563 (0.020) | −0.065 (0.041) | 4.3 | −0.3 | −0.2 | −0.5 | 0.9 | 0.0 | 0.9 |
| Guatemala | 15 | 0.552 (0.020) | 0.027 (0.037) | 4.5 | −0.3 | −0.2 | −0.4 | 1.2 | −0.6 | 0.7 |
| Cuba | 70 | 0.549 (0.017) | −0.024 (0.040) | 4.4 | −1.4 | 0.2 | −1.2 | 1.1 | 1.8 | 2.9 |
| Venezuela | 27 | 0.490 (0.015) | 0.043 (0.047) | 3.5 | 0.4 | −0.3 | 0.1 | −1.0 | 1.1 | 0.0 |
| Colombia | 30 | 0.547 (0.017) | 0.036 (0.037) | 4.3 | −0.5 | 0.1 | −0.4 | 0.8 | −0.3 | 0.5 |
| Brazil | 25 | 0.573 (0.015) | −0.055 (0.038) | 4.3 | −0.7 | 0.0 | −0.8 | 1.0 | 0.0 | 1.0 |
| Cluster A | | | | | | | | | | |
| Ecuador | 21 | 0.496 (0.011) | −0.034 (0.038) | 3.6 | 0.2 | −0.4 | −0.1 | −0.7 | 0.5 | −0.2 |
| Peru | 20 | 0.515 (0.017) | 0.064 (0.033) | 3.9 | 0.0 | 0.2 | 0.3 | 0.0 | −0.4 | −0.4 |
| Bolivia | 30 | 0.527 (0.017) | 0.035 (0.037) | 3.8 | −0.2 | 0.2 | 0.1 | −0.4 | −0.5 | −0.9 |
| Paraguay | 29 | 0.478 (0.015) | 0.026 (0.048) | 3.5 | 0.6 | 0.2 | 0.8 | −1.0 | −0.4 | −1.4 |
| Chile | 20 | 0.473 (0.013) | −0.052 (0.042) | 3.6 | 0.5 | −0.2 | 0.3 | −0.8 | 0.3 | −0.6 |
| Argentina | 25 | 0.498 (0.018) | 0.013 (0.044) | 4.1 | 0.3 | 0.0 | 0.3 | 0.5 | −0.7 | −0.3 |
| Uruguay | 24 | 0.428 (0.024) | 0.146 (0.061) | 3.3 | 1.1 | −0.4 | 0.8 | −1.6 | 1.2 | −0.4 |
| TOTAL | 350 | 0.556 (0.004) | 0.077 (0.011) | 4.5 | | | | | | |

[a] Values obtained via bootstrapping, using 1000 replicates, and sample size fitted to 23 (harmonic mean of the sample sizes of the populations studied).
[b] corresponding to Nei's minimum distance.

between-breeds diversity.

Petit et al. (1998) used Hurlbert's (1971) rarefacted number of alleles per locus ($k$) to assess the contribution of the $i^{th}$ population to total allelic richness as $C_T^g(i) = \frac{\hat{k}_T^g - \hat{k}_{T/i}^g}{k_T^g - 1}$, where $\hat{k}_T^g$ is Hurlbert's (1971) estimator of total allelic richness in the whole analysed population, and $\hat{k}_{T/i}^g$ is the estimator of total allelic richness when the $i^{th}$ population is excluded. The partitioning of $C_T^g(i)$ in two components, i.e.,$C_W^g(i)$, which is the contribution to the total allelic richness due to the own allelic richness of the $i^{th}$ population, and $C_B^g(i)$ which is the contribution due to its divergence, can be obtained as $C_W^g(i) = \frac{1}{n}\left(\frac{\hat{k}_i^g - \hat{k}_{T/i}^g}{\hat{k}_T^g - 1}\right)$ and $C_B^g(i)$ simply by difference $C_B^g(i) = C_T^g(i) - C_W^g(i)$.

Note that positive contributions to diversity from a given population assessed using the method of Caballero and Toro (2002) mean that the remaining dataset increases the overall diversity when the population is removed and, consequently, the assessed population does not contribute significantly to overall diversity (i.e., it would not be preferred for conservation). On the other hand, positive contributions to diversity assessed using the method of Petit et al. (1998) mean that the remaining dataset has a lower number of alleles than the original one and, consequently, the population considered contributes significantly to overall diversity (i.e., it would be preferred for conservation). Examples of the interpretation of such parameters can be found elsewhere (Álvarez et al., 2011, 2010, 2009; Bozzi et al., 2012).

### 2.3. Geographical-based analyses

Estimates of $H_e$, $F_{IS}$, $k_{(24)}$ and within-, between- and total contributions to diversity assessed using both Caballero and Toro (2002) and Petit et al. (1998) methods applied to each donkey population analysed were used to construct interpolation maps drawn using the Inverse Distance Weighted (IDW) option of the Spatial Analyst Extension of ArcView (www.esri.com/software/arcgis/). The IDW (here fitted to a power of two) assumes that each input point has a local influence which diminishes with distance. Interpolation surfaces were divided into eight equal classes. For each country, nodes were fitted roughly averaging the sampling locations (Fig. S1) and also described in Jordana et al. (2016). Linear (flight) distances between nodes were computed using Distance calculator (http://www.distance.to/) in miles and converted into kilometres. The total and within-population contribution to gene diversity as well as total contribution to allelic richness of the various breeds were regressed on their linear kilometric distance relative to the Cuban, Mexican and Venezuelan nodes using Proc Reg of SAS/STAT (SAS Institute, Cary, NC, USA). After applying the Bonferroni correction for multiple tests, each of the three estimated regression coefficients was tested for a statistical significance level of 0.0116 for $\alpha = 0.05$.

## 3. Results

Parameters characterising genetic variability in the analysed dataset and the contributions of each population to diversity are given in Table 1. Note that both the $H_e$ and $F_{IS}$ values presented here depart from those reported in Jordana et al. (2016), due to the adjustment for sample size carried out in the current analysis to avoid bias and to make values fully comparable.

In general, populations classified into Cluster B (northern region) had higher $H_e$ values (from 0.490 ± 0.015 for Venezuela to 0.573 ± 0.015 for Brazil), and allelic richness (ranging from 3.5 for Venezuela to 4.5 for Guatemala) than the populations assigned to Cluster B, in which the Uruguayan donkey had the lowest values (0.428 ± 0.024 for $H_e$ and 3.3 for rarefacted allelic richness). Parameter $F_{IS}$ did not show a consistent pattern, with values ranging from −0.065 ± 0.041 (Mexico) to 0.146 ± 0.061 (Uruguay). These three basic parameters for the assessment of genetic diversity were projected on geographic maps (Fig. 1). From a geographical point of view, $H_e$, $F_{IS}$, and $k_{(24)}$ mainly mirror local genetic events. Gene diversity (Fig. 1A) and even allelic richness (Fig. 1C) did not separate with clarity populations belonging to Cluster A from those included into Cluster B mainly due to the relatively low genetic diversity of the Venezuela population. In these two plots the higher diversity could be identified in Mexico and Guatemala and, to a lesser extent (Fig. 1C), in Colombia, Cuba and Brazil. Fig. 1B, illustrating variation in $F_{IS}$ values, shows that this parameter reflects recent history of the assessed populations.

Contributions to total diversity for each of the thirteen analysed American donkey populations, assessed via Nei's gene diversity and rarefacted allelic richness gave a different picture of the analysed scenario (Table 1). Except for Venezuela, all populations assigned to Cluster B contributed to total diversity whatever the method considered (negative values for $gGD_T$ and positive values for $C_T^{24}$) while those populations classified into Cluster A showed the opposite pattern. For both $gGD_T$

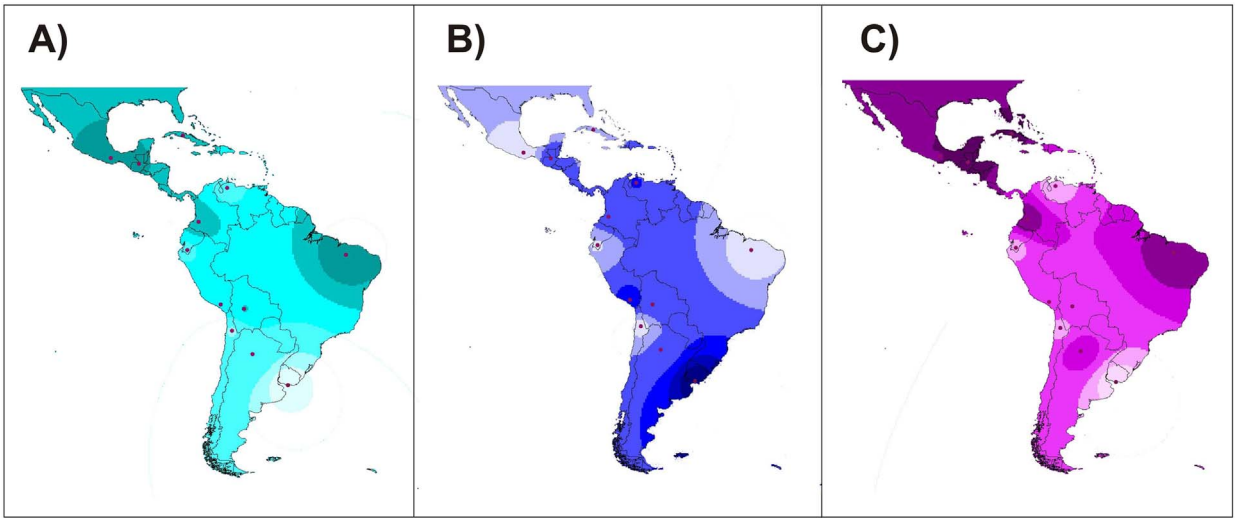



**Fig. 1.** Synthetic maps illustrating geographical variation of expected heterozygosity ($H_e$; Plot A), heterozygote deficiency due to population subdivision ($F_{IS}$; Plot B) and rarefacted number of alleles per locus ($k_{(24)}$; Plot C), in the American donkey populations. The darkest spots illustrate those areas gathering higher (more positive) values for the analysed parameters. Note that in the case of $F_{IS}$, lighter spots mean heterozygote excess while darker spots mean heterozygote deficiency.

(−1.2%) and $C_T^{24}$ (2.9%) the Cuban population had the highest contributions to diversity while the Paraguay population had the lowest contributions to $gGD_T$ (0.9%) and to $C_T^{24}$ (−1.4%). The genetic contribution of the Cuban donkey to the total gene diversity of the American donkey is mainly due to the within-population component ($gGD_W$ = − 1.4%) while the within- and between-population allelic richness, computed according to Petit et al. (1998) are well balanced ($C_W^{24}$ =1.1% and $C_B^{24}$ =1.8%). Projecting of $gGD_T$ and $C_T^{24}$ on geographical maps suggested that the main genetic diversity spot of the species was in Cuba and provided a better differentiation between the areas in which Clusters A and B are located and, mainly in the case of $gGD_T$ (Fig. 2A), a soft cline in contributions to genetic diversity can be detected following a route from Colombia to Uruguay, via Ecuador and Peru. Furthermore, the separation of the two main geographical areas of spreading of the species in America is much clearer at the allelic richness level (Fig. 2B), showing the high homogeneity of the populations belonging to Cluster A for this parameter. Nevertheless, an increased contribution to gene diversity and allelic richness could also be detected in donkeys from northeast Brazil (Figs. 2A and B). As expected, the within- and between-breed components of the contributions to total diversity had lower discriminant power than the full contribution to gene diversity and allelic richness (Fig. S2). When only the between-population component of differentiation was considered ($gGD_B$ and $C_B^{24}$) the Venezuelan and, to a lesser extent, the Uruguayan populations would merit particular attention. Overall, the within- and between-population components of the parameters used to assess contributions to diversity did not depict a clear pattern of genetic variation.

The relationship between the computed parameters and linear kilometric distance to the Cuban, Mexican and Venezuelan nodes, was assessed via regression. In these analyses, the regression on linear kilometric distance to Cuba of total contribution to gene diversity ($gGD_T$; $b = 0.000208$; $P < 0.004$; $R^2 = 0.558$), total contribution to allelic richness ($C_T^{24}$; $b = -0.000370$; $P < 0.006$; $R^2 = 0.521$) and within-populations contributions to gene diversity ($gGD_W$; $b = -0.000205$; $P = 0.012$; $R^2 = 0.453$) were all statistically significant at a level of 0.0116 for $\alpha = 0.05$ (Fig. S3). Regressions on linear kilometric distance to Mexico and Venezuela were not statistically significant (not shown). Even though the Mexico, Guatemala and Venezuela donkey populations showed moderately high genetic differentiation levels relative to the other American donkey populations, regression of genetic parameters on the linear kilometric distance would support the hypothesis considering the Antilles as the putative original breeding nucleus of the species on the American continent.

## 4. Discussion

The main aim of this research was to assess if historical information on the routes of spreading of donkey across the American continent has genetic support. To deal with this task we used genetic parameters ($H_e$ and $k_{(n)}$) which have been shown to be useful in characterising geographical areas acting as the source of genes (the so called “abundant centres” or “hot spots” of variability) and distinguishing them from contact zones (the so called “melting pots”) in which genetic variability results from the influence of various abundant centres (Comps et al., 2001; Eckert et al., 2008; Petit et al., 2003). Nevertheless, our aim has strong concerns. Although important, the number of individuals and populations sampled are limited if a geographical area including the American continent is considered. A larger sample size per population would have increased the statistical reliability of genetic parameter estimates within population, whereas increasing the number of populations sampled would have increased the statistical power to detect geographical patterns of genetic variation. In this scenario, it is not surprising that the genetic signal provided by our dataset was subtle, and that parameters $H_e$ and $k_{(n)}$ may not properly reflect the genetic differences among possible abundant centres and spreading areas (Eckert et al., 2008).

The genetic scenario of the donkey has particular concerns that make it difficult to assess geographical patterns of genetic variation for the species on the American continent. These peculiarities include: a) the American donkey population is strongly structured North-South (Jordana et al., 2016) and some geographical variables, such as latitude or distance, may artificially co-vary with genetic diversity; b) the present genetic variability of donkey populations surrounding the Caribbean area does not present large differences, therefore making it difficult to identify an abundant centre for the species on the American continent (Table 1); and c) many sampled populations suffered recent stochastic local events that are likely to have erased the original genetic signal linked to the spreading of the species. In these circumstances, it is highly advisable to use methods which, in a relative manner, quantify the contributions to diversity, allowing a correct partitioning of gene diversity (Caballero and Toro, 2002) and allelic richness (Petit et al., 1998) among the abundant centre and populations across the range of spreading of the species.

### 4.1. Population structure

Results shown here are fully consistent with the scenarios depicted

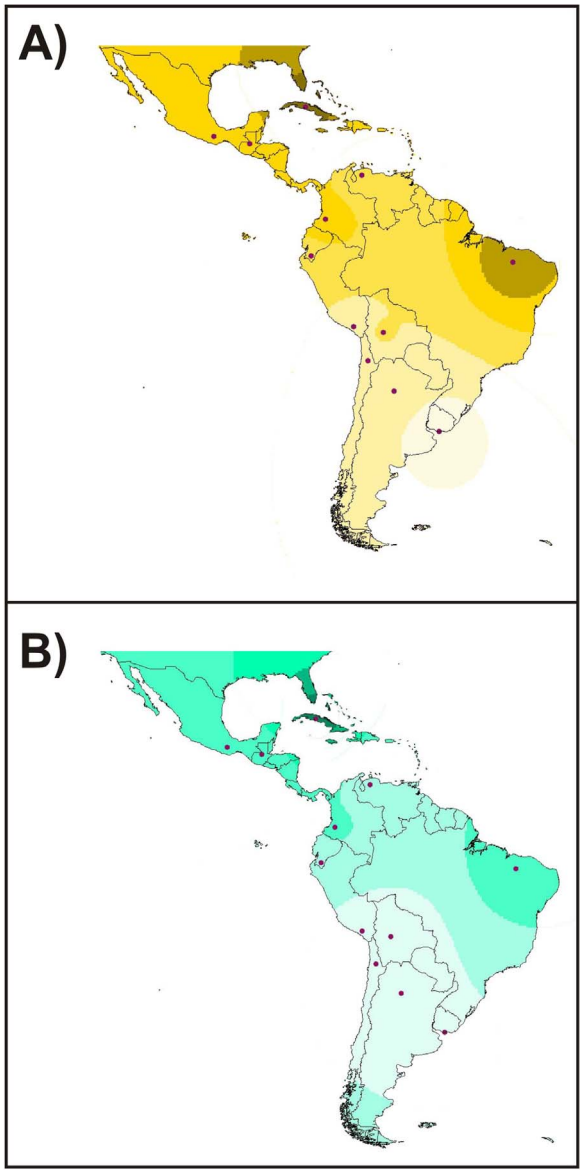



**Fig. 2.** Synthetic maps illustrating geographical variation of the total contribution to gene diversity ($gGD_T$; Plot A) and the total contribution to allelic richness ($C_T^{24}$; Plot B), in the American donkey populations. To facilitate the interpretation of the maps, Caballero and Toro's (2002) contributions to diversity have been multiplied by "−1" and therefore, in all cases the darker spots illustrate those areas gathering higher contributions to diversity.

by Jordana et al. (2016) on American donkey. There is a strong metapopulation substructure among populations classified into either Cluster A or Cluster B. In this respect, current analyses suggest that populations classified into Cluster A are a sample of the genetic background of Cluster B. With no exception, the removal (extinction) of any population belonging to Cluster B leads to a decrease of overall genetic diversity, both at the gene diversity level (negative $gGD_T$ values) and at the allelic richness level (positive $C_T^{24}$ contributions). However, the opposite pattern is found for populations belonging to Cluster A.

To illustrate this issue in a different manner, contributions to diversity at the Cluster level were assessed (Table S1). Cluster B gathered all genetic diversity of the metapopulation. The extinction of the populations belonging to this cluster would decrease the overall American donkey gene diversity by roughly 8% and would affect dramatically the number of alleles in the metapopulation (19.1%). However, the extinction of the donkey populations classified into Cluster A would increase overall gene diversity by 2.2% due to a balance of the allelic frequencies in the remaining population and, in practical terms, would not affect allelic richness (0.4%). In other words, most (if not all) alleles identified in the individuals assigned to Cluster A are present in Cluster B while the opposite is not true.

A quick inspection of these results could suggest that, in the case of continent-wide strategies for the conservation of the species, it would not be necessary to spend effort in the conservation of the southern-most American donkey populations (see Álvarez et al., 2011, 2010, 2009, as examples of the application of the currently tested methods). However, these results must be more likely interpreted in light of the particular genetic scenarios of the populations concerned. Within Cluster A, the population showing the highest deficit of heterozygotes (Uruguay; $F_{IS} = 0.146 \pm 0.061$) and the lowest gene diversity and lowest allelic richness values gave contributions to both $gGD_T$ and $C_T^{24}$ comparable or higher to those of other populations in the cluster (e.g. Paraguay donkey for $gGD_T$ and $C_T^{24}$). This is consistent with a framework of a noticeable genetic identity (Table S2). No matter the different recent history of the populations sampled, Cluster A, as a whole, appears to result from a common founder process with no further relevant gene flow. Considering this, efforts for preservation the genetic variability of this genetic background should be implemented on that Cluster as a whole rather than at the national population level.

### 4.2. Identification of the abundant centre

When dispersal is at random, stochastic processes affecting genetic diversity in a metapopulation are more likely to occur in peripheral populations that can experience subsequent cycles of extinction, recolonization and associated founder events and population bottlenecks, therefore allowing a clear identification of the abundant centre (Eckert et al., 2008). This is not the case of our dataset, therefore making a straightforward identification of an abundant centre of the species on the American continent difficult. A number of populations showing heterozygote deficiency, characterised by positive and high $F_{IS}$ values, can be found in the surroundings of the Caribbean area and in the Southern populations (Table 1; Fig. 1B). The scenarios of, on one hand, the Venezuela population and, on the other, the group formed by the Peru and Bolivia populations are particularly appealing. Although Venezuela is expected to have been quickly colonized by the species after its introduction onto the American continent (Delgado et al., 2010) and that Venezuela donkey individuals have a noticeable genetic identity with Cuban donkey (Jordana et al., 2016), the Venezuela population has $H_e$ and $k_{(24)}$ values fully comparable with those of the populations classified into Cluster A and had negligible contributions to total diversity ($gGD_T = 0.1\%$; $C_T^{24} = 0.0\%$). The Peruvian plateau and surrounding territories were expected to gather a relatively unique genetic diversity, resulting from the well-documented early foundation in this area as a reproductive nucleus of donkeys using feral Jamaican donkeys, which later acted as a major genetic source for the Pampas region (Laguna, 1991; Yanes, 2005). In consequence, it would be expected that the Peruvian and Bolivian donkeys would have higher $k_{(24)}$ and, at least, comparable $H_e$ values relative to their neighbouring populations (Comps et al., 2001). However, this scenario could more likely be identified in the Argentinean donkey population which has a $k_{(24)}$ value comparable to those of the populations belonging to Cluster B. Again, the Peruvian and Bolivian donkey, together with the Chilean population, have probably experienced strong population bottlenecks erasing the genetic signal of the earlier stages of spreading of the species throughout the Continent.

In the current research, the assessment of contributions to total diversity ($gGD_T$ and $C_T^{24}$) appear to be superior to their counterparts' base parameters ($H_e$ and $k_{(24)}$, respectively) in the identification of abundant nucleus. The assessment of geographical patterns of genetic diversity usually assumes that geographically peripheral populations should exhibit lower genetic diversity and higher genetic differentiation than central populations. Considering the within-populations $H_e$ and

$k_{(24)}$ values, some Central American countries (Mexico and Guatemala) and even Brazil could be considered the centre of the species on the American continent (Table 1, Fig. 1A and C). However, this link with Central America is not consistent with the historical information on the introduction of the species onto the American continent (Laguna, 1991; Rodero et al., 1992; Yanes, 2005), while for Brazil there are indications that donkeys were brought directly from Portugal or Cape Verde (McManus et al., 2010). On the other hand, in full agreement with historical sources suggesting that the Greater Antilles were the first breeding nucleus, the contribution to total diversity for both gene diversity and allelic richness of the Cuban population is significantly higher than any other in our dataset (Table 1).

The hypothesis considering the Cuba donkey populations as the more likely representative of the original abundant centre of the species on the Continent was tested via regression of $gGD_T$ and $C_T^{24}$ on linear kilometric distance (Fig. S3). The regressions of such parameters on distance to Mexico or Venezuela were not statistically significant. Moreover, the significant regression of the within-breed contributions to gene diversity ($gGD_w$) on the linear distance to Cuba shows that this donkey population gathers the highest relative number of alleles per locus but also the more balanced within-population allelic frequencies (Fig. S3C). Actually, while allelic richness is expected to be highly reduced by stochastic processes which are more frequent in the fringes of spreading of a species, such processes do not affect $H_e$ to a large extent due to the low influence of rare alleles in $H_e$. In spite of the fact that donkey populations classified into Cluster A have relatively low values of $H_e$ and $k_{(24)}$, as a whole, the Cluster A shows a genetic diversity similar to the cluster B (He=0.582 vs He=0.584, respectively), regardless of the lower allelic richness of Cluster A (5.6 *vs* 6.6, respectively; see Table S1). This result characterises Cluster A as a clear case of single-structured population. Although donkey populations assigned to that Cluster gathered basically the same alleles at any typed loci, they had very unbalanced allelic frequencies, but the merging of the national populations into a single metapopulation (Cluster A) gave a scenario with balanced frequencies and, therefore, with high expected heterozygosity. The contributions to total diversity ($gGD_T$ and $C_T^{24}$) appear useful to reject the existence of a secondary breeding nucleus for the species in Argentina as well. Nevertheless, even though the Argentina donkey population nearly had the highest values for $H_e$ and $k_{(24)}$ within Cluster B, (Table 1), its values for $gGD_T$ and $C_T^{24}$ are essentially the same as those of the Peruvian population (0.3% and −0.3%, respectively), therefore making it not possible to reject the historical documentation on this issue (Laguna, 1991; Yanes, 2005).

### 4.3. Insights of general interest

Jordana et al. (2016) gave the first insights on the main pathways of the colonisation process of the species in America. Our results confirm that Central America donkey populations gather a significant part of the original genetic diversity introduced into America probably being the most likely representatives of the first move of donkey onto the mainland Continent (Delgado et al., 2010; Sponenberg, 1992). Moreover, our results support the suggestion by Jordana et al. (2016) considering the Venezuela donkey population due to a unique founder event with no relevant ensuing gene flow. Actually, values of $gGD_T$ and $C_T^{24}$ estimated for the Venezuela donkey, together with their high genetic identity (see Table S2), would support consideration of the Venezuelan donkey being a Cuban donkey subpopulation.

The role of Mexico, Guatemala and Brazil in the spreading of the species on the American continent is intriguing. Despite the high $H_e$ and $k_{(24)}$ values of those populations, when $gGD_T$ and $C_T^{24}$ are considered, the Mexico, Guatemala donkey populations would be more likely identified as early derivatives of Antillean (here Cuban) donkey. The Brazil donkey population scenario is not so clear. That population had the highest $H_e$ value and $gGD_T$ and $C_T^{24}$ contributions only lower than those of the Cuban donkey (Table 1). Jordana et al. (2016) identified the highest number of private alleles in Brazilian donkey among the American populations of the species, suggesting the existence of a relatively unique genetic background, in the American framework, due to the influence of the Portuguese route during the Conquest or as a result of more recent gene flow (McManus et al., 2010; Primo, 2004). Our results would be consistent with a different genetic history of the Brazilian donkey, but the genetic signal traceable to present day is not sufficiently strong to support the existence of a second, autonomous, centre of diffusion for the species on the American continent.

Furthermore, Jordana et al. (2016) suggested that Colombia could appear as a contact zone between Clusters A and B and that a more recent Iberian (Portuguese) donkey genetic influence on the Brazilian population could not be rejected (Primo, 2004). Regarding the Colombia donkey, the current results could more suggest that the population is representative of the genetic variability on the edge of Cluster B. While $k_{(24)}$ values of a hypothetical Colombian "admixed" population would not be likely to increase due to the poor allelic richness of Cluster A, $H_e$ is expected to increase in a contact zone. However, neither the $H_e$ $(0.547 \pm 0.015)$ nor, particularly, $gGD_T$ (−0.4%) values of the Colombian population fit well to this expectation. With respect to a historically recent influence of Iberian donkey on the Brazilian population, and most other donkey populations in our dataset, our results suggest that, if existent, such processes have not given a clear genetic signal. A more or less recent introgression of either Iberian or African donkey genes into an American population would have dramatically affected its allelic richness and, although our results do not allow the rejection of this hypothesis, this scenario does not fit well with the $k_{(24)}$ and $C_T^{24}$ values computed for the Brazilian donkey (Table 1).

Finally, from the routes of spreading of the species across the Continent summarised by Jordana et al. (2016) our results suggest that the most likely are those considering an early introduction of the species into Central America from the Greater Antilles, with a secondary moving of Antillean donkey into Venezuela and probably Brazil, to be spread via commercial routes until reaching Colombia and, perhaps, Ecuador. The foundation of a breeding nucleus of the species in the Peruvian Plateau from which the current Southern donkey populations of America could have been formed (Laguna, 1991; Yanes, 2005) is the most likely explanation for the formation of the current donkey populations included in Cluster A. The low genetic diversity assessed in the donkey populations near the Plate River prevent accepting the hypothesis suggesting that the Antillean donkey could have directly influenced Southernmost donkey populations via maritime routes and upstream that river (Delgado et al., 2010).

## 5. Conclusions

The genetic characterization of populations and their relative contribution to overall diversity of the species is essential to establish future conservation priorities and strategies. Likewise, the identification of geographical patterns of genetic variation further informs on the development of the species. The ascertainment of such geographical patterns allows checking the consistency of the different historical sources in the spreading of the species in a given area. The projection in geographical maps of genetic parameters such as relative contributions to gene diversity and allelic richness confirmed the Greater Antilles as the entry area of the donkey into the Continent. Furthermore, the Andean route appears as the main route of the spreading of the species into South America. The genetic signal traceable for other historical routes used during the earlier stages of spreading is weak. The results obtained will be useful to support decision making on the conservation strategies contributing to the maintenance of biodiversity.

## Conflict of interest

The authors declare that they have no conflict of interest for this manuscript.

## Acknowledgments

This paper was partially funded by the Ministry of Agriculture, Livestock, Fisheries, Food and the Natural Environment of the Government of Catalonia. We also thank the farmers and the various breeders' associations for their help and collaboration throughout the sampling period. The research was developed in the context of the BIODONKEY consortium of the CONBIAND Network.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.livsci.2016.12.014.

## References

Álvarez, I., Gutiérrez, J.P., Royo, L.J., Fernández, I., Goyache, F., 2009. Quantifying diversity losses due to selection for scrapie resistance in three endangered Spanish sheep breeds using microsatellite information. Prev. Vet. Med. 91, 172–178.
Álvarez, I., Royo, L.J., Pérez-Pardal, L., Fernández, I., Payeras, L., Goyache, F., 2010. Assessing losses of genetic variability in the endangered Mallorquí horse. Czech J. Anim. Sci. 55, 456–462.
Álvarez, I., Royo, L.J., Pérez-Pardal, L., Fernández, I., Lorenzo, L., Goyache, F., 2011. Assessing diversity losses due to selection for coat colour in the endangered bay-Asturcón pony using microsatellites. Livest. Sci. 135, 199–204.
Aranguren-Méndez, J., Jordana, J., Gómez, M., 2001. Genetic diversity donkey breeds using microsatellite DNA markers. Genet. Sel. Evol. 33, 433–442.
Bozzi, R., Álvarez, I., Crovetti, A., Fernández, I., De Petris, D., Goyache, F., 2012. Assessing priorities for conservation in Tuscan cattle breeds using microsatellites. Animal 6, 203–211.
Brookshier, F., 1974. Following the Conquistadors (The burro)Publishing Division of theUniversity. University of Oklahoma Press, Norman, OK, USA, 213–238.
Caballero, A., Toro, M.A., 2002. Analysis of genetic diversity for the management of conserved subdivided populations. Conserv. Genet. 3, 289–299.
Cervantes, I., Pastor, J.M., Gutiérrez, J.P., Goyache, F., Molina, A., 2011. Effective population size as a measure of risk status in rare breeds: the case of three Spanish ruminant breeds. Livest. Sci. 138, 202–206.
Comps, B., Gömöry, D., Letouzey, J., Thiébaut, B., Petit, R.J., 2001. Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. Genetics 157, 389–397.
Delgado, J.V., León, J.M., Gómez, M., Nogales, S., Camacho, M.E., 2010. Las razas ovinas ibéricas y su participación en la colonización de Iberoamérica (Biodiversidad Ovina Iberoamericana -Caracterización y Uso Sustentable)Servicio de Publicaciones de la Universidad de Córdoba. Universidad de Córdoba, Córdoba, Spain, 17–30.
Eckert, C.G., Samis, K.E., Lougheed, S.C., 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Mol. Ecol. 17, 1170–1188.
Gutiérrez, J.P., Royo, L.J., Álvarez, I., Goyache, F., 2005. MolKin v2.0, a computer program for genetic analysis of populations using molecular coancestry information. J. Hered. 96, 718–721.
Hurlbert, S.H., 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52, 577–586.
Jordana, J., Ferrando, A., Miró, J., Goyache, F., Loarca, A., Martínez López, O.R., Canelón, J.L., Stemmer, A., Aguirre, L., Lara, M.A.C., Álvarez, L.A., Llambí, S., Gómez, N., Gama, L.T., Martínez, R.D., Pérez, E., Sierra, A., Contreras, M.A., Guastella, A.M., Marletta, D., Arsenos, G., Curik, I., Landi, V., Martínez, A., Delgado, J.V., 2016. Genetic relationships among American donkey populations, insights into the process of colonization. J. Anim. Breed. Genet. 133, 155–164.
Laguna, E., 1991. El Ganado Español, un descubrimiento para América. Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain.
McManus, C., Paiva, S., Louvandini, H., Melo, C., Seixas, L., 2010. Jumentos no Brasil. INCT: Inf. Genético-Sanitária da Pecuária Bras. (Ser. Técnica: Genética), 1–10.
Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, USA.
Petit, R.J., El Mousadik, A., Pons, O., 1998. Identifying populations for conservation on the basis of genetic markers. Conserv. Biol. 12, 844–855.
Petit, R.J., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Muller-Starck, G.M., Demesure-Musch, B., Palme, A., Martin, J.P., Rendell, S., Vendramin, G.G., 2003. Glacial refugia: Hotspots but not melting pots of genetic diversity. Science 300, 1563–1565.
Primo, A.T., 2004. América: conquista e coloniçao. A fantástica história dos conquistadores ibéricos e seus animais na era dos descobrimentos. Movimento, Porto Alegre, Brazil.
Rodero, A., Delgado, J.V., Rodero, E., 1992. Primitive andalusian livestock and their implication in the discovery of America. Arch. Zootec. 41, 383–400.
Santos, S.A., Sereno, J.R.B., Mazza, M.C.M., Mazza, C.A., 1992. Origin of the Pantaneiro horse in Brazil. Arch. Zootec. 41, 371–381.
Sponenberg, D.P., 1992. Colonial Spanish sheep, goats, hogs, and asses in the United States. Arch. Zootec. 41, 415–419.
Yanes, J.E., 2005. Antecedentes históricos. In: Ministerio de Agricultura. In: Alimentación, Pesca y (Ed.), Razas asnales autóctonas españolas. Ministerio de Agricultura, Pesca y Alimentación, Madrid, Spain, 39–97.

# Contributions to diversity rather than basic measures of genetic diversity characterise the spreading of donkey throughout the American continent

J. Jordana[a,*], F. Goyache[b], A. Ferrando[a], I. Fernández[b], J. Miró[c], A. Loarca[d], O.R. Martínez López[e], J.L. Canelón[f], A. Stemmer[g], L. Aguirre[h], M.A.C. Lara[i], L.A. Álvarez[j], S. Llambí[k], N. Gómez[l], L.T. Gama[m], R.D. Martínez[n], E. Pérez[o], A. Sierra[p], M.A. Contreras[q], V. Landi[r], A. Martínez[r], J.V. Delgado[r]

**Supplementary Table S1** Number of samples (N), within-population expected heterozygosity ($H_e$), deficit of heterozygotes due to population subdivision ($F_{IS}$) rarefacted (to 330 copies) and contributions (in percentage) to within-population, between-population and total diversity of each of the two genetic Clusters identified by Jordana et al. (2016) in American donkey, assessed via Nei's gene diversity and rarefacted allelic richness. Abbreviations mean the following: $gGD_W$: contribution to within-population gene diversity; $gGD_B$: contribution to between-population gene diversity; $gGD_T$: total contribution to gene diversity; $C_W^{330}$: contribution to within-population allelic richness; $C_B^{330}$: contribution to between-population allelic richness; $C_T^{330}$: total contribution to allelic richness. Note that, on the contrary to the methods by Caballero and Toro (2002), the assessment of diversity via rarefacted allelic richness (Petit et al., 1998) does not use the whole remaining dataset as the basis of computations but the individual breed; in consequence, the assessment of the diversity explained by a cluster of populations cannot be straightforwardly done and has been approached by fitting the removed breeds as a single breed assuming the risk of artificially inflating the diversity of the new-composite populations.

| Populations | N | $H_e$[1] | $F_{IS}$[1] | $k_{(330)}$ | $gGD_W$ | $gGD_B$[2] | $gGD_T$ | $C_W^{330}$ | $C_B^{330}$ | $C_T^{330}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Cluster B | 181 | 0.584 (0.006) | 0.059 (0.015) | 6.6 | -5.4 | -2.3 | -7.7 | 9.3 | 9.8 | 19.1 |
| Cluster A | 169 | 0.582 (0.006) | 0.098 (0.016) | 5.6 | 4.4 | -2.3 | 2.2 | -9.3 | 9.8 | 0.4 |

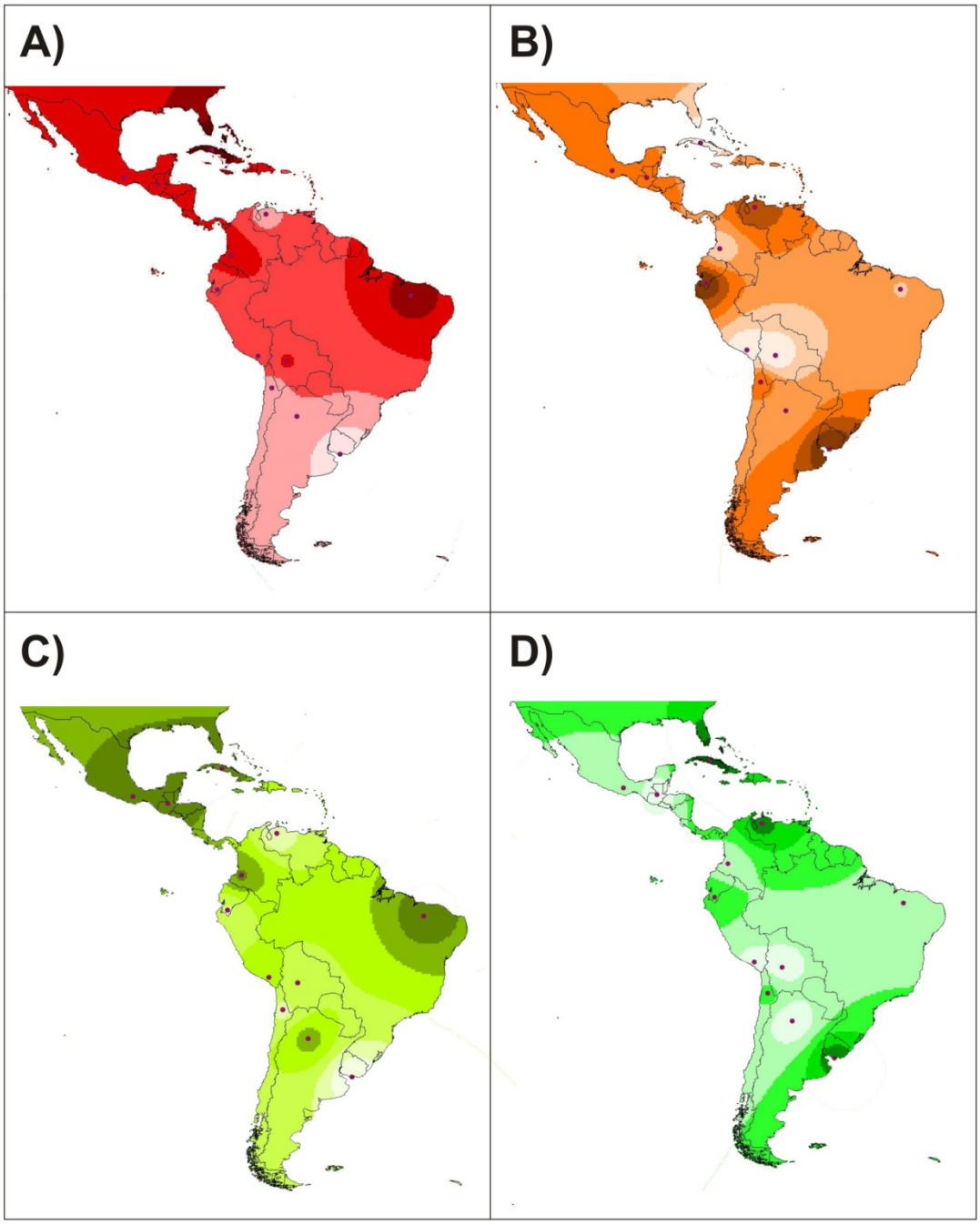
[1]Values obtained via bootstrapping, using 1000 replicates using the corresponding N as sample size.
[2]corresponding to Nei's minimum distance.

**Supplementary Table S2** Between-population molecular coancestry (genetic identity; above diagonal) and Nei's minimum distance (below diagonal) matrices obtained via bootstrapping, using 1000 replicates, and sample size fitted to 23 (harmonic mean of the sample sizes of the populations studied). Standard errors of the estimates are in brackets. The higher genetic identity (molecular coancestry higher than 0.45; in bold) values were assessed among the populations included in Cluster A identified by Jordana et al. (2016). Most Nei's minimum genetic distance values higher than 0.12 (in bold) were found in pairs of populations including the Mexico, Guatemala and Venezuela donkey.

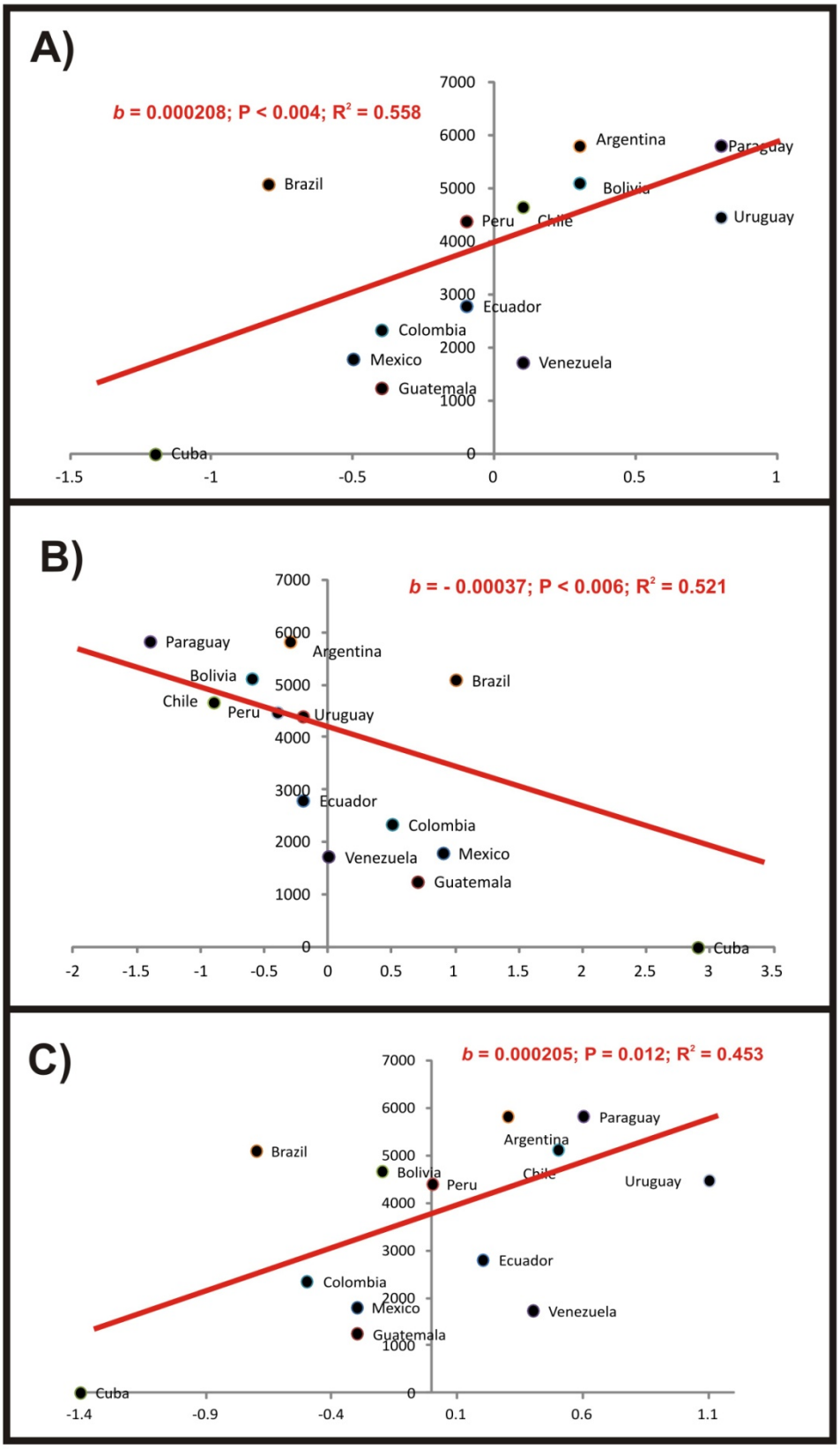
| Country | Mexico | Guatemala | Cuba | Venezuela | Colombia | Brazil | Ecuador | Peru | Bolivia | Paraguay | Chile | Argentina | Uruguay |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [B]Mexico | | 0.407 (0.012) | 0.390 (0.013) | 0.398 (0.015) | 0.373 (0.015) | 0.381 (0.013) | 0.390 (0.011) | 0.403 (0.013) | 0.397 (0.013) | 0.406 (0.014) | 0.398 (0.013) | 0.393 (0.014) | 0.412 (0.014) |
| [B]Guatemala | 0.083 (0.019) | | 0.401 (0.012) | 0.407 (0.013) | 0.386 (0.012) | 0.383 (0.013) | 0.393 (0.011) | 0.411 (0.013) | 0.404 (0.012) | 0.415 (0.012) | 0.396 (0.012) | 0.402 (0.012) | 0.416 (0.012) |
| [B]Cuba | **0.129 (0.027)** | 0.113 (0.018) | | 0.443 (0.012) | 0.413 (0.011) | 0.387 (0.012) | 0.398 (0.012) | 0.419 (0.012) | 0.406 (0.012) | 0.435 (0.013) | 0.413 (0.010) | 0.413 (0.013) | 0.430 (0.013) |
| [B]Venezuela | **0.171 (0.033)** | **0.161 (0.026)** | 0.079 (0.015) | | 0.431 (0.012) | 0.397 (0.012) | 0.415 (0.013) | 0.434 (0.012) | 0.429 (0.013) | **0.450 (0.012**) | 0.424 (0.013) | 0.434 (0.013) | 0.449 (0.012) |
| [B]Colombia | **0.175 (0.036)** | **0.154 (0.024)** | 0.088 (0.018) | 0.108 (0.020) | | 0.393 (0.013) | 0.391 (0.015) | 0.427 (0.015) | 0.418 (0.014) | 0.433 (0.013) | 0.430 (0.015) | 0.427 (0.013) | 0.436 (0.014) |
| [B]Brazil | **0.126 (0.029)** | **0.132 (0.024)** | **0.125 (0.020)** | **0.162 (0.022)** | 0.112 (0.019) | | 0.396 (0.011) | 0.415 (0.012) | 0.402 (0.013) | 0.422 (0.012) | 0.409 (0.012) | 0.407 (0.012) | 0.429 (0.013) |
| [A]Ecuador | **0.185 (0.026)** | 0.082 (0.010) | 0.080 (0.010) | **0.199 (0.024)** | 0.069 (0.009) | 0.069 (0.009) | | **0.451 (0.011)** | 0.436 (0.013) | **0.465 (0.012)** | **0.450 (0.011)** | **0.450 (0.012)** | **0.468 (0.012)** |
| [A]Peru | **0.134 (0.025)** | 0.055 (0.008) | 0.049 (0.007) | **0.138 (0.019)** | 0.040 (0.008) | 0.040 (0.008) | 0.092 (0.018) | | **0.453 (0.013)** | **0.476 (0.012)** | **0.473 (0.012)** | **0.462 (0.012)** | **0.483 (0.012)** |
| [A]Bolivia | **0.137 (0.026)** | 0.056 (0.008) | 0.056 (0.008) | **0.136 (0.019)** | 0.048 (0.008) | 0.048 (0.008) | 0.113 (0.019) | 0.055 (0.014) | | **0.459 (0.012)** | **0.464 (0.013)** | **0.455 (0.014)** | **0.468 (0.013)** |
| [A]Paraguay | **0.163 (0.031)** | 0.069 (0.010) | 0.051 (0.008) | **0.137 (0.021)** | 0.052 (0.009) | 0.052 (0.009) | 0.097 (0.019) | 0.056 (0.014) | 0.079 (0.016) | | **0.473 (0.01)** | **0.475 (0.013)** | **0.507 (0.013)** |
| [A]Chile | **0.187 (0.028)** | 0.092 (0.010) | 0.076 (0.010) | **0.202 (0.025)** | 0.068 (0.011) | 0.068 (0.011) | 0.135 (0.021) | 0.067 (0.015) | 0.073 (0.015) | 0.103 (0.019) | | **0.471 (0.012)** | **0.490 (0.011)** |
| [A]Argentina | **0.176 (0.032)** | 0.073 (0.010) | 0.063 (0.008) | **0.155 (0.021)** | 0.058 (0.009) | 0.058 (0.009) | 0.113 (0.017) | 0.066 (0.0130) | 0.069 (0.016) | 0.075 (0.014) | 0.089 (0.018) | | **0.500 (0.013)** |
| [A]Uruguay | **0.193 (0.029)** | 0.092 (0.014) | 0.081 (0.011) | **0.185 (0.023)** | 0.071 (0.012) | 0.071 (0.012) | 0.137 (0.018) | 0.086 (0.019) | 0.105 (0.018) | 0.075 (0.017) | 0.115 (0.019) | 0.070 (0.016) | |

A and B as superscripts mean that the population is classified into the Clusters A or B, respectively identified by Jordana et al. (2016).


**Supplementary Fig. S1.** Geographical location of sampled animals. Figures within circles correspond to the number of donkeys samples in a given area. Within country, the sampling areas are the following: a) Mexico: 1 Tabasco (4), 2 Oaxaca (10); b) Guatemala: 3 Huehuetenango (15); c) Cuba: 4 Granma (24), 5 Santiago de Cuba (19), 6 Guantanamo (11), 7 Villa Clara (9), 8 Pinar del Río (7); d) Venezuela: 9 Apure (6), 10 Barinas (13), 11 Noroeste (5), 12 Nueva Esparta (3); e) Colombia: 13 Nariño (3), 14 Valle del Cauca (18), 15 Magdalena (9); f) Brazil: 16 Ceará (25); g) Ecuador: 17 El Oro (1), 18 Loja (16), 19 Zamora-Chinchipe (4); h) Peru: 20 Apurimac (9), 21 Puno (9), 22 Cusco (2); i) Bolivia: 23 Cochabamba (31); j) Paraguay: 24 Presidente Hayes (19), 25 Central (3); 26 Alto Paraguay (7); k) Chile: 27 Atacama (20); l) Argentina: 28 Jujuy (5), 29 Salta (6), 30 Catamarca (1), 31 Tucuman (2), 32 Santiago del Estero (4), 33 Cordoba (2), 34 Buenos Aires (5); m) Uruguay: 35 West (4), 36 South (6), 37 East (14).




**Supplementary Fig. S2.** Synthetic maps illustrating geographical variation of the within- ($gGD_W$; Plot A) and between-population ($gGD_B$; Plot B) contributions to gene diversity and the within- ($C_W^{24}$; Plot C) and between-population ($C_B^{24}$; Plot D) contributions to allelic richness, in the American donkey populations. To facilitate the interpretation of the maps, Caballero and Toros's (2002) contributions to diversity have been multiplied by "-1" and therefore, in all cases the darkest spots illustrate those areas gathering higher contribution values for the analysed parameters.




**Supplementary Fig. S3.** Regression of total contribution to gene diversity ($gGD_T$; Plot A), total contribution to allelic richness ($C_T^{24}$; Plot B) and within-population contribution to gene diversity ($gGD_W$; Plot C) on linear kilometric distance (on Y-axis) between each sampling donkey node and the Cuban node.