Received: 1 July 2016 | Accepted: 21 December 2016
DOI: 10.1111/jbg.12255

WILEY Journal of Animal Breeding and Genetics

ORIGINAL ARTICLE

The legacy of Columbus in American horse populations assessed by microsatellite markers

O. Cortés[1] | **S. Dunner**[1] | **L.T. Gama**[2] | **A.M. Martínez**[3] | **J.V. Delgado**[3] | **C. Ginja**[4] | **L.M. Jiménez**[5] | **J. Jordana**[6] | **C. Luis**[7,8,9] | **M.M. Oom**[10] | **D.P. Sponenberg**[11] | **P. Zaragoza**[12] | **Biohorse Consortium** | **J.L. Vega-Pla**[13]

[1]Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense de Madrid, Madrid, Spain
[2]CIISA, Faculdade de Medicina Veterinária, Universidade de Lisboa, Lisboa, Portugal
[3]Departamento de Genética, Universidad de Córdoba, Córdoba, Spain
[4]Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-InBIO), Universidade do Porto, Vairão, Portugal
[5]Departamento de Producción Animal, Facultad de Medicina Veterinaria y de Zootecnia, Universidad Nacional de Colombia Sede Bogotá, Bogota, Columbia
[6]Departamento de Ciencia Animal y de los Alimentos, Facultad de Veterinaria, Bellaterra, Spain
[7]Centro Interuniversitário de História das Ciências e da Tecnologia (CIUHCT), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
[8]Museu Nacional de História Natural e da Ciência (MUHNAC), Universidade de Lisboa, Lisboa, Portugal
[9]CIES-UL, Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
[10]cE3c – Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
[11]Virginia-Maryland College of Veterinary Medicine, Virginia Tech, Blacksburg, VA, USA
[12]Laboratorio de Genética Bioquímica, Facultad de Veterinaria, Universidad de Zaragoza, Zaragoza, Spain
[13]Laboratorio de Investigación Aplicada, Cria Caballar de las Fuerzas Armadas, Cordoba, Spain

Correspondence
Ó. Cortés, Departamento de Producción Animal, Facultad de Veterinaria, Madrid, Spain
Email: ocortes@vet.ucm.es

Funding information
Fundação para a Ciência e a Tecnologia, Grant/Award Number: IF/00866/2014.

Summary

Criollo horse populations descend from horses brought from the Iberian Peninsula over the period of colonization (15th to 17th century). They are spread throughout the Americas and have potentially undergone genetic hybridization with other breeds in the recent past. In this study, 25 autosomal microsatellites were genotyped in 50 horse breeds representing Criollo populations from 12 American countries (27 breeds), breeds from the Iberian Peninsula (19), one breed each from France and Morocco and two cosmopolitan horse breeds (Thoroughbred and Arabian). The genetic relationships among breeds identified five clusters: Celtic; Iberian; North American with Thoroughbred influence; most Colombian breeds; and nearly all other Criollo breeds. The group of "all other Criollo breeds" had the closest genetic relationship with breeds originating from the Iberian Peninsula, specifically with the Celtic group. For the whole set of Criollo breeds analysed, the estimated genetic contribution from other breeds was approximately 50%, 30% and 20% for the Celtic, Iberian and Arab-Thoroughbred groups, respectively. The spatial distribution of genetic diversity indicates that hotspots of genetic diversity are observed in populations from Colombia, Ecuador, Brazil, Paraguay and western United States, possibly indicating points of

---

Biohorse Consortium: The list of Biohorse Consortium members are given in the Acknowledgment section (http://biohorse.jimdo.com).

arrival and dispersion of Criollo horses in the American continent. These results indicate that Criollo breeds share a common ancestry, but that each breed has its own identity.

KEYWORDS
criollo breeds, genetic relationships, genetic variability, microsatellite markers

1 | INTRODUCTION

Domestic horses were first brought to America from the Iberian Peninsula during the second voyage of Columbus in 1493 (Laguna, 1991; Rodero, Rodero, & Delgado, 1992). Following arrival to the Caribbean islands, horses rapidly disseminated throughout the American continent following two major routes. One to Mexico (1519) and from there to New Mexico and California, and another to Colombia and Venezuela (1509), from where horses then spread to other South American regions (Laguna, 1991; Rodero et al., 1992). Throughout the 16th century, other groups of horses continued to arrive from the Iberian Peninsula to the West Indies, the Rio de la Plata region and Brazil (Primo, 2004;).

The direct descendants of horses brought to America since the arrival of Columbus are generally known as Criollos and are found in most American countries. Criollo populations each have their own identity, but in some cases, there is some gene flow among them, especially in populations located in adjacent areas with no geographical barriers. Also, in some parts of the United States, the introduction of cosmopolitan European breeds and the development of more intensive production and breeding systems since the 19th century have favoured the formation of new admixed breeds that replaced many of the original Criollo populations.

Microsatellite markers have been widely used to analyse the genetic diversity and population structure of some Iberian and Criollo horse populations (Cañón et al., 2000; Cothran, Canelon, Luis, Conant, & Juras, 2011; Cothran & Luis, 2005; Luís, Juras, Oom, & Cothran, 2007; Marletta et al., 2006; Morais, Oom, Malta-Vacas, & Luis, 2005; Vega-Pla, Calderón, Rodríguez-Gallardo, Martinez, & Rico, 2006; Vega-Pla et al., 2005). However, a comprehensive analysis of Criollo and Iberian horse breeds using autosomal markers has not yet been reported. In this study, we investigated the genetic diversity of Criollo and Iberian horse breeds to obtain clearer understanding of their current status and provide new information regarding breed relationships, genetic structure and breed influences in Criollo horse populations.

2 | MATERIALS AND METHODS

A group of 50 horse breeds from Spain (16), Portugal (3), France (1), Morocco (1), America (27) and two cosmopolitan breeds (Arabian and Thoroughbred) were included in the study (Figure S1).

The 2,385 samples analysed came from 16 countries (Europe, 3; Africa, 1; America, 12) as summarized in Table 1.

Genomic DNA was extracted from hair or blood samples. A total of 25 autosomal microsatellite loci recommended by FAO/ISAG for genetic diversity studies in horses were genotyped as follows: VHL20, HTG4, AHT4, HMS7, ASB2, ASB17, HMS6, AHT5, ASB23, HTG10, HMS3, LEX33, TKY287, TKY294, TKY297, TKY301, TKY312, TKY321, TKY325, TKY333, TKY337, TKY341, TKY343, TKY344 and TKY394.

Various software packages were used for different statistical analyses, as summarized in Table S1. Mean number of alleles (MNA) per locus and population, and unbiased expected (He) and observed (Ho) heterozygosities were computed with the Microsatellite Toolkit, while *F*-statistics and their 95% confidence intervals over all loci after 1,000 bootstraps and F_{ST} genetic distances among populations were obtained with the FSTAT software. An exact test for Hardy–Weinberg (HW) equilibrium across loci and populations was performed using the Markov chain method to estimate *p*-values, as implemented in GENEPOP 1.2. A neighbour-net dendrogram was constructed with the SPLITSTREE4 package based on the matrix of F_{ST} genetic distances among breeds. Factorial correspondence analysis (FCA) was performed using GENETIX v.4.04, and the STRUCTURE v.2.1 software was used to investigate the genetic structure of the 50 horse populations and to estimate the proportions of individual genotypes derived from the inferred clusters. The DISTRUCT v.1.1 software was used to obtain a graphical display of individual membership coefficients in each ancestral population, considering the run with the highest posterior probability of the data at each K value ($K = 2$ to $K = 5$).

A maximum likelihood estimation of admixture proportions was carried out to estimate the genetic contributions from various breed sources to the formation of Criollo breeds, using the LEADMIX software. The potential source breeds from Portugal, Spain and Cosmopolitan breeds were grouped into three independent groups according to the results achieved with STRUCTURE. The Criollo breeds were grouped in seven clusters, according to their geographical location, except for Peruano de Paso and Cr. Salvador,

TABLE 1 Breed, acronyms, country of origin and genetic descriptive parameters of the samples analysed. Sample size (N), observed (Ho) and expected (He) heterozygosities, mean number of alleles per locus (MNA) and F_{IS} (within-breed inbreeding coefficient) with correspondence level of significance (*$p < .01$) values

	Breed	Acronym	Country	*N*	Ho	He	MNA	F_{IS}
America	Criollo Argentino	CAR	Argentina	19	0.66	0.70	5.8	0.063
	Mangalarga Brasil	MGB	Brazil	30	0.74	0.76	7.0	0.035
	Marajoara Brasil	MJB	Brazil	54	0.77	0.79	8.0	0.03*
	Pantaneiro Brasil	PNT	Brazil	60	0.79	0.78	7.6	−0.018
	Puruca	PUR	Brazil	47	0.77	0.78	8.1	0.011
	Chilote	CHI	Chile	63	0.73	0.78	8.7	0.061*
	Isla De Pascua	ISP	Chile	27	0.74	0.75	6.4	0.009
	Criollo Colombiano Vaqueria	CVC	Colombia	29	0.82	0.79	7.8	−0.03
	Paso Fino Colombiano	CPF	Colombia	72	0.73	0.74	8.2	0.012
	Trocha Pura Colombiana	CTC	Colombia	72	0.74	0.75	8.3	0.014
	Trocha Y Galope Colombiano	CTP	Colombia	69	0.74	0.75	8.5	0.005
	Trote Y Galope Colombiano	CTG	Colombia	72	0.74	0.75	8.3	0.015
	Cubano	CUB	Cuba	23	0.71	0.69	5.5	−0.025
	Criollo De Ecuador	CEC	Ecuador	43	0.80	0.80	8.5	0.006
	Criollo Del Salvador	CSA	El Salvador	20	0.72	0.75	7.3	0.042
	Panameño	PAN	Panama	30	0.76	0.80	8.1	0.046*
	Criollo Paraguayo	CPA	Paraguay	42	0.77	0.80	8.7	0.034*
	Peruano De Paso	PEP	Peru	50	0.76	0.76	7.7	0.002
	Criollo Uruguayo	CUR	Uruguay	33	0.67	0.75	7.3	0.1*
	Appaloosa	APA	USA	44	0.76	0.80	8.8	0.053*
	Morgan	MOR	USA	87	0.74	0.76	8.1	0.02
	Quarter Horse	QHO	USA	95	0.77	0.77	8.5	0.008
	Rocky Mountain	RMU	USA	58	0.75	0.72	7.8	−0.034
	Saddlebred	SDB	USA	94	0.69	0.70	7.1	0.013
	Spanish Mustang	SMU	USA	48	0.75	0.77	8.0	0.019
	Tennessee Walking	TWH	USA	80	0.73	0.73	8.0	−0.003
	Mount Taylor Mustang	MTM	USA	53	0.75	0.77	8.0	0.03
Iberian Peninsula								
Celtic type	Merens Pirineo Frances[a]	MER	France	16	0.67	0.71	5.8	0.079
	Garrano	GAR	Portugal	25	0.74	0.77	7.2	0.035
	Asturcon	AST	Spain	31	0.70	0.73	6.8	0.04
	Burguete	BUR	Spain	50	0.72	0.76	8.1	0.053*
	Cabalo do Monte Galego	GAL	Spain	40	0.73	0.75	7.8	0.02
	Hispano-Breton	BRT	Spain	30	0.76	0.77	7.5	0.007
	Jaca Navarra	JNV	Spain	28	0.74	0.77	8.0	0.032
	Losino	LOS	Spain	58	0.76	0.75	8.0	−0.011
	Mallorquin	MAL	Spain	48	0.76	0.74	7.2	−0.02
	Menorquin	MEN	Spain	57	0.73	0.75	7.8	0.017

(Continues)

TABLE 1 (Continued)

	Breed	Acronym	Country	N	Ho	He	MNA	F_{IS}
	Monchino	MON	Spain	30	0.76	0.81	7.8	0.057*
	Pirinenc Català	PIC	Spain	41	0.69	0.74	7.8	0.071*
	Potoka	POT	Spain	24	0.77	0.77	7.5	−0.002
Iberian type	Lusitano	LUS	Portugal	21	0.70	0.75	6.4	0.071*
	Sorraia	SOR	Portugal	23	0.62	0.64	4.5	0.036
	Hispano Árabe	HIA	Spain	40	0.75	0.77	7.2	0.019
	Marismeño	MAR	Spain	63	0.73	0.76	7.9	0.033*
	Pura Raza Española	PRE	Spain	60	0.71	0.74	7.6	0.034*
	Retuertas	RET	Spain	67	0.74	0.71	6.8	−0.031
	Trotador Español	TRO	Spain	46	0.74	0.76	7.1	0.021
Cosmopolitan Breeds	Arabian	ARA	60	0.65	0.68	6.3	0.031	Arabian
	Thoroughbred	PSI	60	0.72	0.73	5.7	0.024	Thoroughbred
Africa	Barb	BER	Morocco	21	0.77	0.76	7.3	−0.008
	Total	50	16	2,385	0.74 ± 0.0143	0.75 ± 0.0174	7.5 ± 1.7	0.02

[a] Located in the boundary between Spain and France.

which were included with Paso Fino Colombiano lineages due to the results of the F_{ST} genetic distances dendrogram. The seven geographical clusters were as follows: CRA – Mangalarga, Marajoara, Pantaneiro, Puruca; CRB – Cr. Colombiano Vaquería, Paso Fino Colombiano, Trocha Pura Colombiano, Trocha y Galope Colombiano, Trote y Galope Colombiano, Peruano de Paso and Cr. Salvador; CRC – Cr. Argentino, Cr. Paraguayo and Cr. Uruguayo; CRD – Isla de Pascua and Chilote; CRE – Appaloosa, Quarter Horse, Morgan, Rocky Mountain, Saddlebred, Spanish Mustang, Tennessee Walking, Mount Taylor Mustang; CRF – Cr. Panameño and Cr. Ecuatoriano; CRG – Cr. Cubano.

The expected heterozygosity estimates for the Criollo horse populations were interpolated on a map depicting the geographical distribution of the various breeds using an Inverse Weighted Distance (IWD) in ARCGIS 9.3 (ESRI 2011). The IWD interpolation was used to predict genetic diversity values outside of a breeds geographical location, assuming that expected heterozygosity values will be more similar among nearby points. A similar approach was used to plot the spatial distribution of F_{IS} estimates for the various Criollo populations, to assess the regional incidence of population bottlenecks.

3 | RESULTS

The microsatellite markers showed a mean number of alleles per locus of 7.5, ranging from 4.5 in Sorraia to 8.8 in Appaloosa (Table 1). Observed and unbiased expected heterozygosities per breed ranged from 0.62 (Sorraia) to 0.82 (Cr. Colombiano Vaqueria) and from 0.64 (Sorraia) to 0.81 (Monchino), respectively, with slight differences among overall means for observed (0.74) and expected (0.75) heterozygosity (Table 1). The highest F_{IS} was found in Cr. Uruguayo (0.100), and a total of 15 horse breeds showed a significant deficit in heterozygosity (see Table 1). The amount of differentiation (F_{ST}) among the 50 populations studied was 0.071 (95% CI 0.066–0.075), while global F_{IS} was 0.017 (95% CI 0.01–0.023).

The F_{ST} genetic distances among breeds are presented in Figure S2. The mean F_{ST} genetic distance for each breed ranged from 0.04 (Cr. Paraguayo and Cr. Ecuatoriano) to 0.15 (Sorraia). The mean F_{ST} genetic distances were higher among Iberian Peninsula horse breeds than among Criollo populations (0.06 and 0.08, respectively; results not shown).

When assessed at lower K-values, the Bayesian cluster analysis implemented by STRUCTURE indicates a subdivision of the horse populations that preceded breed formation (Rosenberg et al., 2002). As shown in Figure S3, results for $K = 2$ indicate a separation of breeds that have a well-known Thoroughbred influence, as is the case of the United States horse breeds and the Trotador Español. Setting $K = 3$ further indicated a new cluster that grouped the Paso Fino Colombiano lineages (Paso Fino Colombiano, Trocha Pura Colombiana, Trocha y Galope Colombiano, Trote y Galope Colombiano). For $K = 4$, the Iberian-type breeds were grouped in a new cluster which also included the Garrano, while Criollo and Celtic-type breeds remained together in the same cluster. Finally, Criollo breeds (except the ones from the United States and Paso Fino Colombiano lineages) and

Iberian Peninsula Celtic populations were further subdivided in two separate clusters when $K = 5$. It is notable that the Spanish Mustang showed stronger signatures of admixture with Iberian Peninsula breeds than the remaining USA horse breeds.

The FCA (Figure S4) and the F_{ST} genetic distance dendrogram (Figure 1) supported the subdivision of USA breeds, except Spanish Mustang, and Paso Fino Colombiano lineages. Furthermore, the FCA analysis showed that the thoroughbred could be differentiated in one separated cluster within the USA group. Both analyses suggested a certain degree of grouping of Iberian and Celtic horse breeds and an intermediate position between these two groups of the majority of Criollo horse populations. Also, it is notable that one Celtic-type (Asturcon) and two Iberian-type (Mallorquin and Menorquin) breeds showed a marked tendency to group together with the Criollo horse breeds (Figures S3, S4 and Figure 1). On the other hand, the Spanish Mustang breed showed a higher genetic similarity to Criollo horse populations from other regions than to those from the United States breeds. The Sorraia is very distant from the centre of the dendrogram in Figure 1, possibly due to its high inbreeding and low genetic diversity.

A total of three clusters for Iberian Peninsula and cosmopolitan breeds were identified with the Bayesian analysis implemented by STRUCTURE. The first group of source breeds included the Iberian-type breeds plus the Garrano (Iberian group), the second group included Celtic breeds (Celtic group) and the third group integrated two well-known cosmopolitan related horse breeds, that is Arabian and Thoroughbred (Arab group). The average contribution of the three ancestral populations to the whole group of Criollo breeds analysed was 48% from the Celtic group, 29% from the Iberian group and 23% from the Cosmopolitan breeds group (Figure 2). Four of the Criollo clusters (CRA, CRB, CRC and CRD) showed a higher

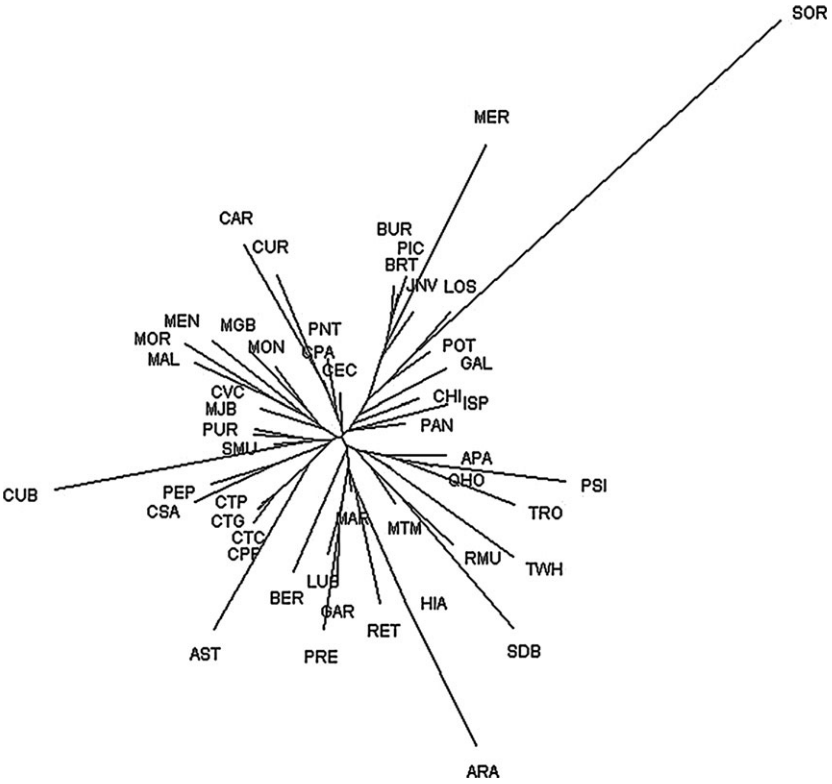

FIGURE 1 The neighbour-joining dendrogram based on F_{ST} genetic distances between horse breeds analysed. Cr. Argentino (CAR), Mangalarga (MGB), Marajoara (MJB), Pantaneiro (PNT), Puruca (PUR), Chilote (CHI), Isla de Pascua (ISP), Cr. Colombiano Vaqueria (CVC), Paso Fino Colombiano (CPF), Trocha Pura Colombiana (CTC), Trocha y Galope Colombiano (CTP), Trote y Galope Colombiano (CTG), Cr. Cubano (CUB), Cr. Ecuatoriano (CEC), Cr. Salvadoreño (CSA), Cr. Panameño (PAN), Cr. Paraguayo (CPA), Peruano de Paso (PEP), Cr. Uruguayo (CUR), Appaloosa (APA), Morgan (MOR), Quarter Horse (QHO), Rocky Mountain (RMU), Saddlebred (SDB), Spanish Mustang (SMU), Tennessee Walking (TWH), Mount Taylor Mustang (MTM), Arabian (ARA), Asturcon (AST), Burguete (BUR), Cabalo do Monte Galego (GAL), Hispano-Breton (BRT), Hispano-Arabe (HIA), Jaca Navarra (JNV), Losino (LOS), Mallorquin (MAL), Menorquin (MEN), Marismeño (MAR), Monchino (MON), Pirinenc Català (PIC), Potoka (POT), Pura Raza Español (PRE), Thoroughbred (PSI), Retuertas (RET), Trotador Español (TRO), Garrano (GAR), Lusitano (LUS), Sorraia (SOR), Merens Pirineo Frances (MER) and Barb (BER)

genetic influence of the Celtic ancestral population, which had a contribution ranging between 43% and 56% to their genetic pool. On the other hand, the Iberian ancestral population had the highest genetic contribution to the CRF cluster (Criollo populations from Panama and Ecuador). As expected, horse breeds in the CRE cluster (breeds from the United States) showed the highest genetic contribution from the Cosmopolitan breeds cluster. The CRG group (Criollo Cubano) had a strong influence from the Celtic cluster, but no influence from the Iberian cluster.

The spatial distribution of expected heterozygosity and F_{IS} values in Criollo horse populations (Figure 3) indicated that hotspots of genetic diversity were observed in populations from Colombia, Ecuador, Brazil, Paraguay and western United States. A reduction in genetic diversity was visible in the Criollo populations from Argentina and Uruguay and, to a lesser extent, the eastern United States, while the remaining Criollo populations sampled show intermediate levels of genetic diversity. The highest deficit in heterozygosity was observed in horses from Uruguay and Argentina and in the Appaloosa breed.

4 | DISCUSSION

The Iberian Peninsula is a hotspot of horse genetic diversity, because it served as a refuge for wild horses in the early and mid-Holocene (Warmuth et al., 2011), so higher levels of horse genetic variability are expected. Our results indicate that American horse populations, even endangered ones, have levels of genetic variability similar to those in Iberian breeds and slightly higher than those reported for other European horse breeds (Barcaccia et al., 2013; Leroy et al., 2009; Thirstrup, Pertoldi, & Loeschcke, 2008). These results are similar to those reported for other Iberian (Cañón et al., 2000; Gómez, Azor, Alonso, Jordana, & Valera, 2012; Luís et al., 2007; Marletta et al., 2006; Morais et al., 2005; Vega-Pla et al., 2005) and Criollo horse populations (Conant, Juras, & Cothran, 2012; Cothran & Luis, 2005; Cothran et al., 2011) in studies using microsatellite markers. Our results indicate a high degree of conservation of the original genetic variability from Iberian breeds in Criollo horse populations. It is also possible that some of these breeds may have received admixed contributions, which could lead to their high levels of genetic

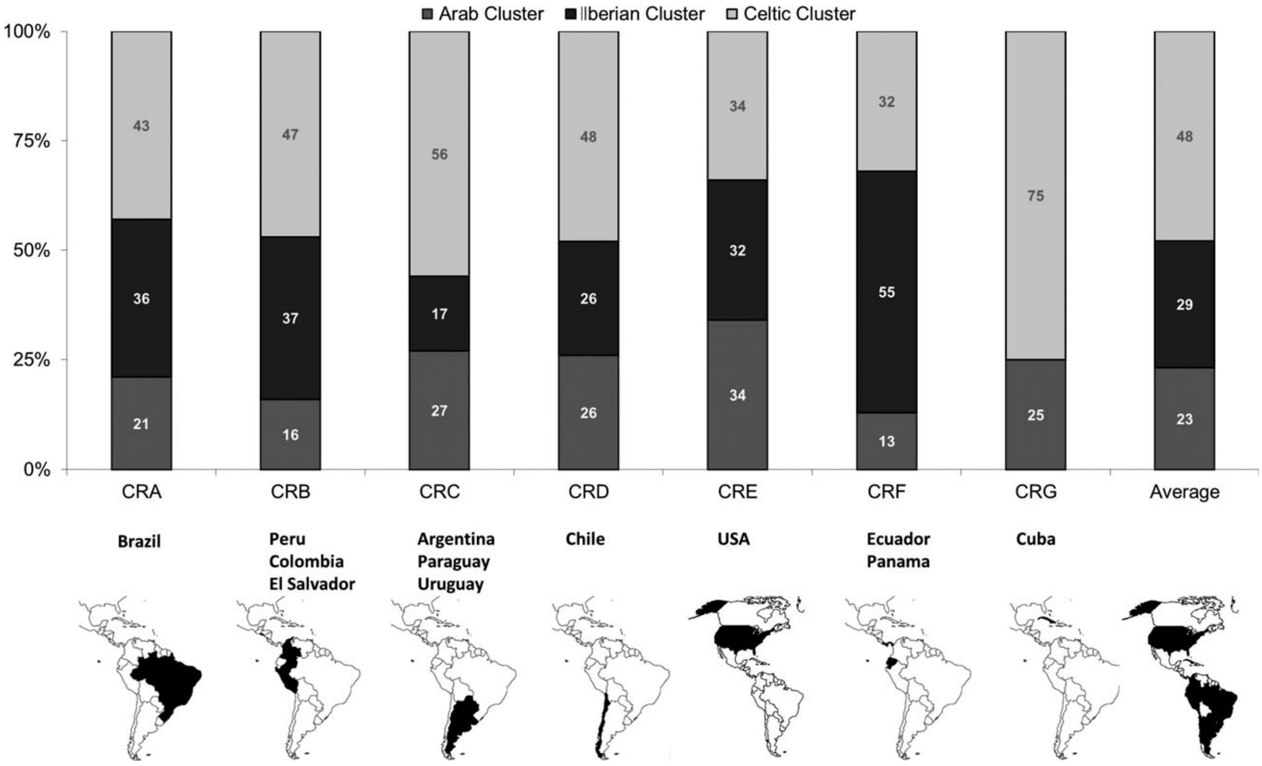

FIGURE 2 Proportional genetic contribution from three ancestral populations of Iberian Peninsula horse breeds to Criollo breeds grouped in eight clusters. Iberian Peninsula clusters: ARAB: Thoroughbred and Arabian; IBERIAN: Lusitano, Pura Raza Español, Marismeño, Garrano, Retuertas, Sorraia and Berberisco; CELTIC: Asturcón, Jaca Navarra, Cabalo do Monte Galego, Monchino, Hispano-Bretón, Losino, Pirinenc Catala, Pottoka and Burguete. CRA: Mangalarga, Marajaora, Pantaneiro, Puruca; CRB: Cr. Colombiano Vaqueria, Paso Fino Colombiano, Trocha Pura Colombiano, Trocha y Galpoe Colombiano, Trote y Galope Colombiano, Peruano de Paso and Cr. Salvadoreño; CRC: Cr. Argentino, Cr. Paraguayo and Cr. Uruguayo; CRD: Isla de Pascua and Chilote; CRE: Appaloosa, Quarter Horse, Morgan, Rocky Mountain, Saddlebred, Spanish Mustang, Tennessee Walking, Mount Taylor Mustang; CRF: Cr. Panameño and Cr. Ecuatoriano; CRG: Cr. Cubano

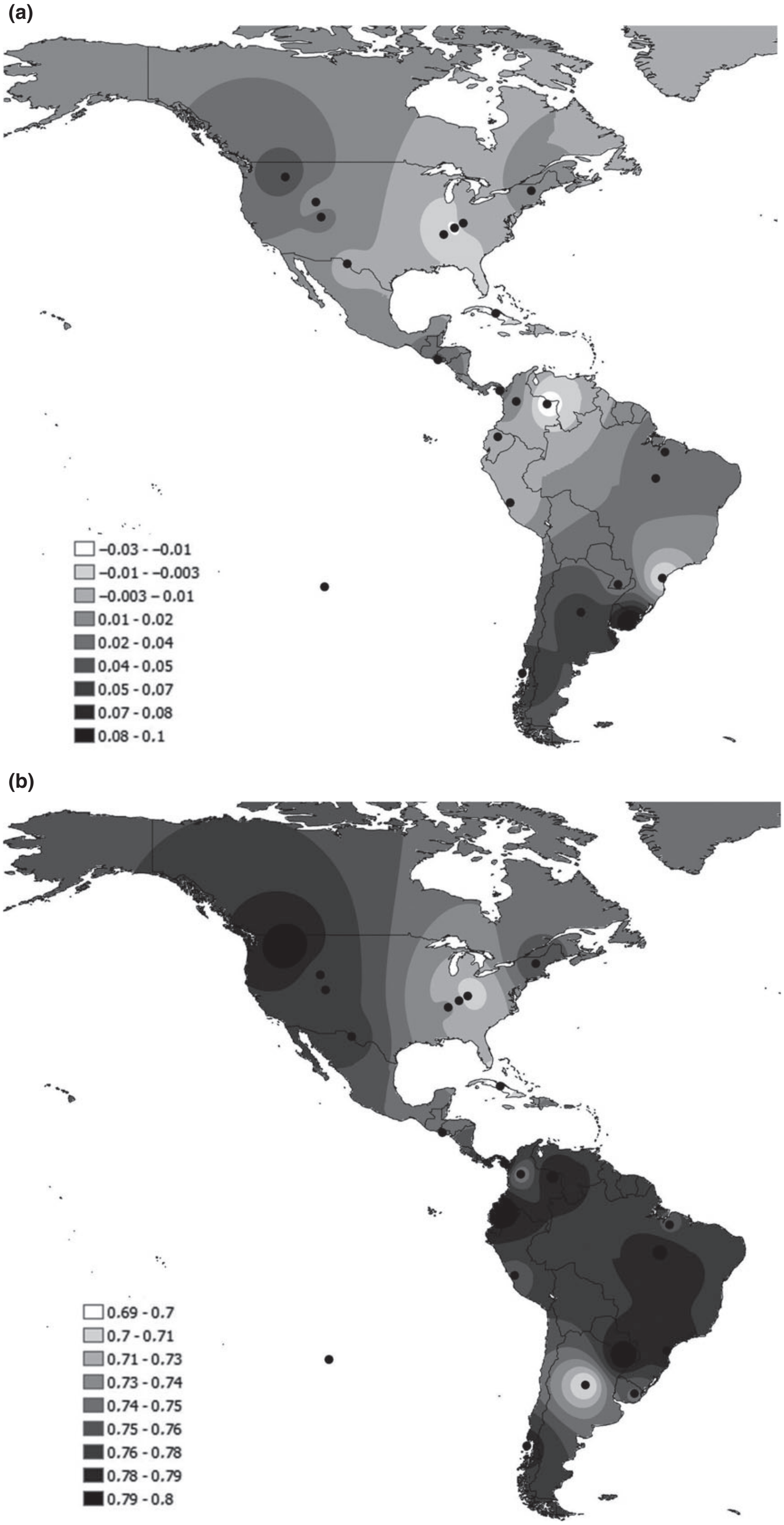

FIGURE 3 The intensity of the grey colour represents the F_{IS} values (a) and the expected heterozygosities (b) obtained with an inverse distance-weighted (IDW) interpolation method using the ArcGIS Spatial Analyst extension. Grey dots represent the geographical location of the American horse populations analysed

variability. It is well known that a bottleneck has occurred in some Criollo horse populations (e.g. Cr. Argentino and Cr. Uruguayo) as a consequence of efforts to recover the ancestral populations from a narrow base in the early 20th century, after most were extensively crossed with stallions from the United Kingdom and France. This bottleneck could be the reason for their comparatively low heterozygosity (<0.7).

Our results reveal a clear subdivision of two major clusters, that is USA breeds, except Spanish Mustang and the Paso Fino Colombiano lineages.

Criollo horse populations in North America were originally established from horses introduced from Santo Domingo to Mexico by Diego Velázquez and Hernán Cortés starting in 1511 (Primo, 2004). These animals were later taken to New Mexico and California and eventually originated feral populations that were sometimes mixed with different European breeds in order to improve their original aptitudes (Clutton-Brock, 1992). For example, the Thoroughbred contributed to the Quarter Horse to increase its speed (Petersen, Mickelson, Cleary, & McCue, 2014), and also to the Rocky Mountain, Tennessee Walking and Saddlebred breeds, which originated in the same geographical area (Dutson, 2005). Thoroughbred, Arabian, Quarter Horse, Morgan and Standardbred influences have been documented in the Appaloosa (Appaloosa Horse Club), while all Morgan horses can trace back to a single stallion, *figure*, whose sire was probably a Thoroughbred (Appaloosa Horse Club). The Mount Taylor Mustang is considered to have a Colonial Spanish phenotype, but also has some variation in coat colours that is not typical of most Iberian horse breeds, such as roan or palomino. Probably, the feral management of Mount Taylor Mustang populations for centuries could have permitted cross-breeding with other geographically close American horse populations, which would justify our results. The Spanish Mustang is considered a direct descendant of Iberian horses with less Thoroughbred influence, which is confirmed by our results and revealed a closer genetic relationship with other Criollo and Iberian populations, rather than with the other North American horse breeds analysed, which were more admixed.

The Paso Fino Colombiano lineages were grouped in a single cluster. The arrival of horses to Colombia followed two main routes, the first coming from Panama and arriving to the northern part of the Andes and the second coming from Venezuela and reaching the plains south of the Andes. The marshy areas of the north favoured the development of a horse with gaits more suitable for agriculture or riding (Jiménez, & Cortés et al., 2012). Years of selection produced the Colombian Paso Fino that was later divided into four lineages based on natural gait ability (CPF, CTC, CTP and CTG). Other breeds such as Peruano de Paso and Cr. Salvadoreño were selected for the same gaits as the Colombian Paso Fino, and some degree of gene flow among them may have occurred, resulting in their location in the same branch of F_{ST} genetic distances matrix dendrogram and the results obtained in STRUCTURE. In contrast, horses raised in the plains to the south of the Andes are called Cr. Colombiano Vaqueria and are mainly used in agricultural activities, but they are mostly isolated in a state of semi-freedom in herds without much selective pressure. This could justify their closer genetic relationship with the Criollo breeds rather than with Paso Fino Colombiano lineages.

Our results suggested the existence of another three major clusters, that is Iberian, Celtic and Criollo populations, the latter with different levels of admixture among them, as expected. Breeds from the Iberian Peninsula divided into two major clusters, Iberian and Celtic, in accordance with their traditional classification. The Garrano was the only exception to this pattern, because it is generally considered to belong to the Celtic group but clustered with the Iberian-type horses, this may reflect recent admixture with breeds of larger size. The hardiness and small size of Celtic-type horses may have favoured their selection for the long journey to the American continent in the initial colonization period, and this influence is still detectable in Criollo populations. Among Celtic breeds, the Asturcon had a stronger influence in the development of Criollo horse populations, especially in Colombian lineages, as revealed by STRUCTURE results at low K-values and the dendrogram of F_{ST} genetic distances. The major dispersion force for the Criollo populations has been genetic drift and adaptation to different environmental conditions since their arrival to the American continent, with some degree of gene flow among them, according to their geographical proximity. The distinction of these breeds is mostly based on morphological differences and/or geographical location, so that different names were assigned to populations which are highly related, but located in different administrative areas. The arrival of European horse breeds during the 19th century could mask the original genetic relationships among Criollo breeds or produce misleading results.

The Cr. Argentino, Cr. Uruguayo, Cr. Paraguayo and Pantaneiro were grouped together within a F_{ST}-based dendrogram and represent a main colonization route through the southern part of the continent. Horses were initially taken to the Pampas region in the mid-16th century by Pedro de Mendoza and later on by Cabeza de Vaca, who brought horses from Santa Catarina (Brazil) to Paraguay, and from there to Argentina and Uruguay (Santos, Sereno, Mazza, & Mazza, 1992), which would justify the close relationship among these populations revealed in our study.

In addition to the Rio de la Plata route, horses were also brought to Brazil in the mid-16th century, either directly from Portugal or through Cape Verde, arriving in the regions corresponding to northeast Brazil and São Paulo (Primo, 2004) and leading to the various Criollo populations presently existing in Brazil. In the particular case of the Marajoara breed, it is believed that it derives from Portuguese horses brought from Cape Verde, which were introduced into the Marajó archipelago at the beginning of the 18th century (Figueiró, Costa, Coelho, Costa, & Marques, 2012). At the end of the 19th century, Shetland ponies imported from France were crossed with Marajoara horses and originated the Puruca breed (Sávio, Evonnildo, Silva, & Schneider, 2008). In our analyses, the Bayesian approach placed the remaining Brazilian horse breeds in addition to Pantaneiro (Puruca, Mangalarga and Marajoara) in the same cluster as the Criollos from Argentina and Uruguay and suggested the existence of some degree of gene flow among them. In the F_{ST} dendrogram, the Puruca and Marajoara Brazilian breeds showed a close genetic relationship, confirming their admixture.

The horses from Mexico were desirable in Cuba during the colonial period, and the presence in Cuba of horses from Mexico is noted in the 19th century, which is reflected by the genetic proximity observed in the neighbour-joining tree (Figure 1) among Cr. Cubano and the Spanish Mustang, which originated from horses initially brought to Mexico in 1,511 by Diego Velázquez (Diaz del Castillo, 2012).

The Chilote and Isla de Pascua horse breeds from Chile were grouped in the same cluster as Iberian Celtic breeds. This was probably as a consequence of the horses introduced in Chile during the 16th century being of Celtic type rather than of the Iberian type, due to their hardiness and small size, which were essential in mountainous regions (Randall, 2008).

The analysis of genetic contributions to Criollos from various breed sources supports the high genetic contribution of the Celtic parental population to Criollo breeds. Their small body size and lower nutritional requirements were important for the long journey across the Atlantic, and their morphological characteristics provided adaptation to the different environmental conditions in the Americas. A higher genetic contribution of the Iberian parental population was detected in group CRF (Cr. Ecuatoriano and Cr. Panameño), as a consequence of the predominant influence in their development of horses from the South of the Iberian Peninsula (Hendrickson, 2013), where the Lusitano and Pura Raza Español are found. As expected, the United States horse breeds revealed a higher contribution of the Arabian/Thoroughbred parental population (34%), even though North American horse breeds still show an important contribution of Iberian Peninsula horse breeds, as shown by the genetic contribution of Celtic (34%) and Iberian (32%) parental populations. In our analyses, the inclusion in the North American cluster of the Spanish Mustang and Mount Taylor Mustang breeds, which are considered to be close descendants of horses brought to America from the Iberian Peninsula, could increase the Iberian and Celtic genetic contribution to the North American cluster, which would not be as important in other breeds from the United States. The null introgression from Iberian parental populations observed in the CRG group (Cr. Cubano) is notable, which probably results from isolation and loss of genetic variability by genetic drift.

The spatial distribution of horse genetic diversity in South America revealed by our study is consistent with the known colonization routes followed since the first arrival of horses to the American continent in 1493, but could also reflect some other influences that have occurred since then. Our results support the existence of three major axes corresponding to areas of high genetic diversity in Criollo horse populations, that is western United States, Colombia-Ecuador and Brazil-Paraguay. These correspond to the major points of arrival of horses to the American continent and could thus represent vestiges of the original colonization routes. Lower levels of genetic diversity are observed in Argentina and the eastern United States, while the highest levels of heterozygosity deficit are observed in Uruguay, Argentina and Chile. The low genetic diversity and high F_{IS} detected in this region could be a consequence of a bottleneck occurring in southern Criollo horse populations at the beginning of the 20th century, when the near extinction of these populations resulted in a notable decrease in genetic diversity in this geographical area.

In conclusion, the majority of the Criollo horse breeds analysed in our study show clear signs of the influence of breeds originating from the Iberian Peninsula, even though other breeds may have also been introduced in more recent years. With the exception of some breeds from the United States that were crossed with Arabian-derived breeds such as the Thoroughbred, the Celtic-type horse populations from the Iberian Peninsula had the greatest genetic influence in the development of Criollo horse populations, followed by Iberian-type horses. Globally, three genetic clusters are well differentiated among Criollo populations: (i) the majority of the USA breeds, (ii) Colombian Paso Fino lineages and (iii) the remaining Criollo breeds, with some degree of genetic substructure among them.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions. We thank the different breeders associations and research groups who kindly provided biological samples used in this study. Members of CONBIAND networks are thanked for valuable cooperation over the years. Catarina Ginja received funding from Fundação para a Ciência e a Tecnologia, Portugal, through an IF2014 Contract Grant (IF/00866/2014).

Members of the Biohorse Consortium: The following members of the Biohorse Consortium have contributed to his study: Lenin Aguirre Riofrío, Universidad Nacional de Loja, Ecuador. Claudia Brandariz, Juan Calderón, Estación Biológica de Doñana, España. Javier Cañón Ferreras, Natalia Sevane, Universidad Complutense de Madrid, España.

María Rosa Costa, José Ribamar Marques, Embrapa Amazônia Oriental, Belen, Pará, Brasil. Carlos Dominguez Sanchez, Universidad Veracruzana, México. Sarah Farman, Founder and Director of AHHC Spanish Mustangs. Ainhoa Ferrando, Universitat Autònoma de Barcelona, España. Mayra Gómez Carpio, Vincenzo Landi, Universidad de Córdoba, España. Inmaculada Martín-Burrriel, Clemetina Rodellar, Universidad de Zaragoza, España. Fernando Mújica, Universidad Austral de Chile, Chile. Lorena Núñez, Universidad Nacional de Asunción, Paraguay. M. Cecilia T. Penedo, University of California, USA. Fabiana Tavares Pires de Souza Sereno, Brasil.

REFERENCES

Appaloosa Horse Club. *Appaloosa Horse Club handbook 2016*. Retrieved from http://www.appaloosa.com/pdfs/rulebook16.pdf

Barcaccia, G., Felicetti, M., Galla, G., Capomaccio, S., Cappelli, K., Albertini, E., … Verini, A. (2013). Molecular analysis of genetic diversity, population structure and inbreeding level of the Italian Lipizzan horse. *Livestock Science*, *151*, 124–133.

Cañón, J., Checa, M. L., Carleos, C., Vega-Pla, J. L., Vallejo, M., & Dunner, S. (2000). The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. *Animal Genetics*, *31*, 39–48.

Clutton-Brock, J. (1992). *Horse power: A history of the horse and the donkey in human societies*. London: Natural History Museum Publications.

Conant, E. K., Juras, R., & Cothran, E. G. (2012). A microsatellite analysis of five Colonial Spanish horse populations of the southeastern United States. *Animal Genetics*, *4*, 53–62.

Cothran, E. G., Canelon, J. L., Luis, C., Conant, E., & Juras, J. (2011). Genetic analysis of the Venezuelan Criollo horse. *Genetics and Molecular Research*, *10*, 2394–2403.

Cothran, E. G., & Luis, C. (2005) Genetic distance as a tool in conservation of rare horse breeds. In I. Bodo, L. Alderson & B. Langolis (Eds.), *Conservation genetics of endangered horse breeds* (pp. 55–71), EAAP 116. Wageningen: Wageningen Academic Publishers.

Diaz del Castillo, A. (2012). *The true history of the conquest of New Spain*. Indianapolis, IN: Hackett Publishing Company, Inc.

Dutson, J. (2005). *Storey's illustrated guide to 96 horse breeds of North America* (pp. 214–216). North Adams, MA: Storey Publishing.

ESRI (2011). *ArcGIS desktop: Release 10*. Redlands, CA: Environmental Systems Research Institute.

Figueiró, M., Costa, J. S., Coelho, L., Costa, M. R., & Marques, J. R. F. (2012). *Equinos (Equus caballus)* da raçz Marajoara em conservaçao na Ilha de Marajó. II Congresso Brasileiro de Recursos Genéticos. Belém, Pará, Brazil: https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/950511/1/594.pdf.

Gómez, M. D., Azor, J. P., Alonso, M. E., Jordana, J., & Valera, M. (2012). Morphological and genetic characterization of Spanish heavy horse breeds: Implications for their conservation. *Livestock Science*, *144*, 57–66.

Hendrickson, S. L. (2013). A genome wide study of genetic adaptation to high altitude in feral Andean Horses of the páramo. *BMC Evolutionary Biology*, *13*, 273.

Jiménez, L. M., Méndez, S., Dunner, S., Cañón, J., & Cortés, O. (2012). Colombian Creole horse breeds: Same origin but different diversity. *Genetics and Molecular Biology*, *35*, 790–796.

Laguna, E. (1991). *El Ganado Español, un descubrimiento para América*. Madrid, Spain: Ministerio de Agricultura, Pesca y Alimentación,.

Leroy, G., Callède, L., Verrier, E., Mériaux, J. C., Ricard, A., Danchin-Burge, C., & Rognon, X. (2009). Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genetics Selection Evolution*, *41*, 5.

Luís, C., Juras, R., Oom, M. M., & Cothran, E. G. (2007). Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. *Animal Genetics*, *38*, 20–27.

Marletta, D., Tupac-Yupanqui, I., Bordonaro, S., García, D., Guastella, A. M., Criscione, A., … Dunner, S. (2006). Analysis of genetic diversity and the determination of relationships among western Mediterranean horse breeds using microsatellite markers. *Journal of Animal Breeding and Genetics*, *123*, 315–325.

Morais, J., Oom, M. M., Malta-Vacas, J., & Luis, C. (2005). Genetic structure of an endangered Portuguese semiferal pony breed, the Garrano. *Biochemical Genetics*, *43*, 347–364.

Petersen, J. L., Mickelson, J. R., Cleary, K. D., & McCue, M. E. (2014). The American Quarter Horse: Population structure and relationship with the Thoroughbred. *Journal of Heredity*, *105*, 148–162.

Primo, A. T. (2004). *América: Conquista e colonizaçao: A fantástica história dos conquistadores ibéricos e seus animais na era dos descobrimentos*. Porto Alegre, Brazil: Movimento.

Randall, R. A. (2008) The Chilean horse. America's oldest registered cowhorse. Chapter I-3. *The untold origins of the South American horses*. Retrieved from http://www.ChileanHorse.com.

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

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic structure of human populations. *Science*, *298*, 2381–2385.

Santos, S. A., Sereno, J. R. B., Mazza, M. C. M., & Mazza, C. A. (1992). Origin of the Pantaneiro horse in Brazil. *Archivos de Zootecnia*, *41*, 371–381.

Sávio, P. R., Evonnildo, C. G., Silva, A., & Schneider, M. P. C. (2008). Genetic variability and efficiency of DNA microsatellite markers for paternity testing in horse breeds from the Brazilian Marajó archipelago. *Genetics and Molecular Biology*, *31*, 68–72.

Thirstrup, J. P., Pertoldi, C., & Loeschcke, V. (2008). Genetic analysis, breed assignment and conservation priorities of three native Danish horse breeds. *Animal Genetics*, *39*, 395–405.

Vega-Pla, J. L., Calderón, J., Rodríguez-Gallardo, P. P., Alcaide, B., Sereno, F., Costa, M. R., … Rico, C. (2005) The Retuertas Horse: The 'Missing Link' in the iberoamerican horse breeds origin? In I. Bodo, L. Alderson & B. Langlois (Eds.), *Conservation Genetics of Endangered Horse Breeds* (pp. 167–176). Wageningen: Wageningen Academic Publishers, EAAP Publication 116.

Vega-Pla, J. L., Calderón, J., Rodríguez-Gallardo, P. P., Martinez, A. M., & Rico, C. (2006). Saving Feral Horse Populations: Does it really matter? A case study of wild horses from Doñana National Park in Southern Spain. *Animal Genetics*, *37*, 571–578.

Warmuth, V., Eriksson, A., Bower, M. A., Cañón, J., Cothran, G., Distl, O., . . . Manica, A. (2011). European domestic horses originated in two holocene refugia. *PLoS One*, *6*, e18194.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Cortés O, Dunner S, Gama LT, et al. The legacy of Columbus in American horse populations assessed by microsatellite markers. *J Anim Breed Genet*. 2017;00:1–11.