The Origins of Iberian Horses Assessed via Mitochondrial DNA

L. J. Royo, I. Álvarez, A. Beja-Pereira, A. Molina, I. Fernández, J. Jordana, E. Gómez, J. P. Gutiérrez, and F. Goyache

From Área de Genética y Reproducción Animal, SERIDA, E-33203 Gijón, Asturias, Spain (Royo, Álvarez, Fernández, Gómez, Goyache); CIBIO-UP, Faculdade de Ciências, Universidade do Porto, Campus Agrário de Vairão, 4485–661 Vairão (VCD), Portugal (Beja-Pereira); Departamento de Genética, Universidad de Córdoba, E-14071-Córdoba, Spain (Molina); Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, E-08193-Bellaterra, Barcelona, Spain (Jordana); Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense de Madrid, E-28040, Madrid, Spain (Gutiérrez).

Address correspondence to Félix Goyache at the address above, or e-mail: fgoyache@serida.org.

Abstract

Despite a number of recent studies that have focused on the origin of domestic horses, genetic relationships between major geographical clusters still remain poorly understood. In this study we analyzed a 296 bp mtDNA fragment from the HVI region of 171 horses representing 11 native Iberian, Barb, and Exmoor breeds to assess the maternal phylogeography of Iberian horses. The mtDNA haplogroup with a CCG motif (nucleotide position 15,494 to 15,496) was the most frequent in Iberian and Barb breeds (0.42 and 0.57, respectively), regardless of geographic location or group of breeds. This finding supports the close genetic relationship between the ancestral mare populations of the Iberian Peninsula and Northern Africa. Phenotypic differences among the Northern and Southern Iberian groups of breeds are not explained by population subdivision based on maternal lineages. Our results also suggest that Northern Iberian ponies—which are phenotypically close to British ponies, especially Exmoor—are the result of an introgression rather than population replacement.

The Iberian Peninsula is an interesting territory for testing different hypotheses on the origin and spread of domestic animals. It has an intermediate geographical position between two continents (Europe and Africa), and it is peripherally located to areas in which the major human and livestock expansions in the Neolithic period are assumed to have occurred (Central-Eastern Europe and the Near and Middle East in Asia).

Iberian horses have been classically divided into two native groups: Celtic ponies and Iberian horses (Aparicio 1944). The first group, mainly settled along the Cantabrian Range in the Northern Atlantic Iberian areas (Figure 1), is subdivided into breeds or populations such as the Garrano in Portugal (which is called Faco galego on the Spanish side of the border) or the Asturcón and Pottoka in Spain. Most of the Northern Iberian breeds are in high risk of extinction and are conserved in environmentally protected rural areas of Spain and Portugal. The *Celtic* designation of the Northern Iberian ponies is roughly based on historical evidence, as the distribution of these breeds coincides with the Iberian areas in which the Celts (Indo-Europeans originating from Central Europe) settled after 1,500 BCE (Villar 1996). The second group is located in the Southern Iberian Peninsula

and includes the two lines of the Spanish purebred horse (Andalusian and Carthusian) and Marismeño (an ancient semiferal horse population located in Doñana Natural Park) in Spain and the Lusitano and the Sorraia in Portugal. Both Northern and Southern Iberian horses can be well differentiated by means of morphological traits (Jordana et al. 1995; Jordana and Parés 1999). Several studies have demonstrated that Northern Iberian strains of horses are phenotypically close to British ponies, especially the Exmoor breed (Álvarez Llana 1995; García-Dory 1980; Martínez Sáiz et al. 1996), suggesting a possible common origin with these breeds (Aparicio 1944; Sotillo and Serrano 1985). The Southern Iberian horses are assumed to have a common origin with the Barb horse on the other side of the Strait of Gibraltar (Aparicio 1944; Jansen et al. 2002; Sotillo and Serrano 1985).

Mitochondrial DNA (mtDNA) has been extensively used to assess the phylogeography and domestication of the horse. Vilà et al. (2001) reported that domestic horses originated from multiple matrilines. Additionally, Jansen et al. (2002) reported extensive genetic diversity summarized in 93 different mtDNA types grouped into 17 distinct phylogenetic clusters. The examination of horse population diversity at the mtDNA level has revealed a consistent absence of geographical
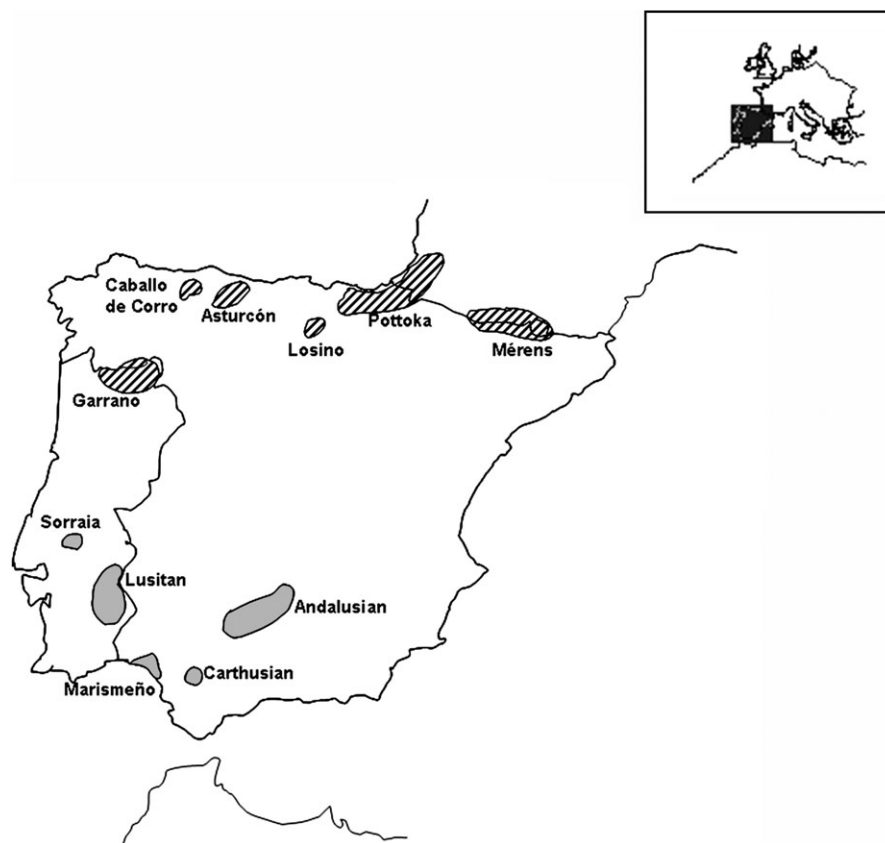
Figure 1. Main geographical location of the Iberian horse breeds sampled in the present study. The Northern Iberian breeds (with bars) include Asturcón, Caballo de Corro, Garrano, Losino, Mérens, and Pottoka. Southern Iberian breeds (in gray) include Andalusian, Carthusian, Lusitano, Marismeño, and Sorraia.

structure (Hill et al. 2002). Only a few of the phylogenetic clusters reported by Jansen et al. (2002) corresponded to groups of breeds or geographic areas; interestingly enough, those groups were the Northern European ponies and the Southern Iberian and Northwestern African breeds.

The aim of this study is to contribute to a comprehensive molecular analysis of the different native strains of horses of the Iberian Peninsula to increase knowledge of the origin and expansion of horses. Iberian horse breeds are compared with major representatives of the Northern European and Northwestern African clusters of horses—namely, Exmoor and Barb horse breeds, which have close phenotypical, geographical, and historical relationships with Northern Iberian and Southern Iberian horse breeds, respectively. Throughout the article, we address the following questions: Is the Iberian population of horses heterogeneous? How are Iberian horses genetically related to major representatives of the Northwestern African and Northern European clusters of horses?

Material and Methods

DNA Extraction, Amplification, and Sequencing

Total DNA was isolated from blood samples following standard procedures (Sambrook et al. 1989). The mtDNA control region was amplified by using the following PCR primers: 5′-ACCATCAACACCCAAAGCTG-3′ and 5′-CCTTGAA-GAAAGAACCAGATG-3′ (which amplify a 361 bp region). The polymerase chain reaction (PCR) protocol consisted of a 20 μl reaction mix containing approximately 50 ng total horse DNA, 0.5 μM each primer, 2.5 mM MgCl₂, 0.2mM dNTPs, and 0.5 U *Taq* polymerase (Biotools, Madrid). The PCR consisted of an initial denaturation step at 94°C for 5 min followed by 35 cycles of 30 s at 94°C, 30 s at 60°C, and 1 min at 72°C, with a final elongation step of 10 min at 72°C. PCR products were purified using ExoSAP-IT (USB, Cleveland, OH) following the manufacturer's recommendations. DNA sequencing was performed from the PCR product in both directions on an ALFexpressII automated sequencer (Amersham Biosciences, Barcelona) using the Thermo Sequenase Cy5 Dye Terminator Cycle Sequencing Kit (Amersham Biosciences, Barcelona).

Samples

Blood samples were obtained from 100 unrelated horse belonging to nine Iberian breeds. The breeds and numbers of individuals (in parentheses) sampled were as follows: Garrano (10) in Portugal, Andalusian (7), Asturcón (12), Caballo de Corro Asturiano (14), Carthusian (10), Losino

(10), Marismeño (12) and Pottoka (13) in Spain, and Mérens (12) in Andorra. The GenBank accession numbers and references for the sequences obtained in the present analysis are from AY519871 to AY519970. To avoid overweighting our analysis with Northern Iberian populations, we gathered mtDNA control region sequences from individuals belonging to the Andalusian (14), Lusitano (16), and Sorraia (7) breeds from the Genbank database (accession numbers AJ413615–28, AJ413736–45, AY246242–47, AF516502–4, and AY246259–65). Figure 1 shows the geographical location of the sampled breeds and how these were grouped into Northern Iberian and Southern Iberian breeds. To ascertain the genetic relationships between Iberian horses and Northwestern African and Northern European clusters (Jansen et al. 2002), we obtained the same mtDNA information from Genbank (accession numbers AJ413682–93, AY246219–24, AF072992–3, and AJ413658–71) of individuals of the Barb (14) and the Exmoor (20) breeds. All together, mtDNA sequences were collected for 138 Iberian horses and a total of 171 samples.

Sequence Analysis

Sequences were aligned using the PILEUP program (Feng and Doolittle 1987) of the Winsconsin Package 10.2 (Genetics Computer Group, Madison, WI) and were checked by eye. Analyses were restricted to 296 bp (nucleotide positions from 15,480 to 15,776) for the purpose of comparing our sequences with those from published data. The nucleotide position (np) numbering follows that of Xu and Arnason (1994; accession number X79547).

Phylogenetic Analysis

Interbreed differentiation was assessed by computing the p distance, which is merely the proportion of nucleotide sites at which the two sequences compared are different, and the linearized Φ_{ST} (Slatkin 1995) values for each population pair using the Arlequin 2.0 software package (Schneider et al. 2000). The statistical significance of the values was estimated by permutation analysis using 1,000 replications.

To identify biologically meaningful, empirically supported phylogenies, maximum parsimony networks were constructed using both the reduced median network method (Bandelt et al. 1995) and the median-joining network method (Bandelt et al. 1999) as implemented in the program NETWORK 3.1.1.1 (available at http://www.fluxus-engineering.com).

Results

The 171 sequences gave 55 different haplotypes (Table 1) with 53 variable and 36 parsimony informative sites. Overall nucleotide diversity was 7.1 (6.9 and 7.1 for the Northern and Southern Iberian groups, respectively). Among the Iberian samples, 45 different haplotytpes were found, 12 of these being shared between the two groups. Average haplotype diversity was 32.4%, ranging from 28.8% in the Southern Iberian

group to 52.1% in the Northern Iberian group. Among the identified haplotypes, the one denominated H1 was the most frequent (24% for the whole data set and 27% for the Iberian samples), followed by H19 (7%). In sum, 29 haplotypes are unique; the Northern Iberian and Southern Iberian groups respectively share 5 and 4 haplotypes with Barb horses, and the Exmoor breed shares 2 haplotypes with the Northern Iberian breeds but does not share haplotypes with the Southern Iberian and Barb breeds. Barb haplotypes are present in a total of 61 Iberian samples, whereas Exmoor haplotypes (H20 and H53) are present in only two Losino and Mérens samples. The most frequent haplotype in the Exmoor breed (H16) is not present in the other breeds from the data set.

Figure 2 shows a median-joining network constructed from the most parsimonious trees supported by the whole data set. The robustness of this network was confirmed by a reduced median network algorithm (not shown). Because of the uneven pattern of substitution in the D-loop region (Sbisà et al. 1997) leading to homoplasies, the network was contracted for character conflicts (Forster et al. 2001). However, a high number of reticulations remain, showing the presence of parallel mutations. Five major clusters (I–V) can be identified. However, there is an absence of a phylogenetic structure in our data set, thus preventing the identification of an evolutionary relationship between clusters. The most striking feature of Figure 2 is the clear differentiation existing between those haplotypes with a CCG motif from np 15,494 to 15,496 (cluster Ia) and the others, except for those included in cluster Ib. Clusters with a TCA motif in the same positions (clusters II–V) are in turn clearly disconnected. Cluster IV includes Sorraias, and cluster IIIa includes most Exmoors. Our results are in agreement with those reported by Jansen et al. (2002). Cluster Ia includes the clusters D1, D2, and D3 as defined by Jansen et al. (2002). Cluster D1 was reported by these authors as being characteristic of the geographical cluster formed by Iberian (Andalusian and Lusitano) and North African horses (Barbs). This family of haplotypes includes 57.1% of the samples in the Barb breed, 45.5% in the Southern Iberian group (notice their absence in Sorraias), and 40.8% in the Northern Iberian group.

Table 2 shows genetic distances between breeds assessed by means of p distance and linearized Φ_{ST} values between breeds. Neither of the distances indicates a clear genetic differentiation between breeds, except for Sorraias. The average p distance and standard error for the whole data set is 0.0251 ± 0.005, and most paired p distances lie within this range. Linearized Φ_{ST} values are in general small, except for most Sorraia pairs. Most Φ_{ST} pairs including Andalusian, Asturcón, Barb, Carthusian, and Garrano breeds are low. Paired distances affecting the Exmoor breed tend to be higher. Some geographically and genetically close breed pairs, such as Asturcón–Caballo de Corro and Andalusian-Carthusian show higher values of differentiation than those aforementioned. Carthusians are a small population of inbred individuals that are the genetic basis of the present Andalusian (Valera et al. in press) and Lusitano horses (d'Andrade 1954). In addition, there is solid historical evidence that

Table 1. Variable sites, cluster assignment, frequency, and number (*k*) of haplotypes by breed, Iberian group of breeds, and totals

Variable sites and nucleotide positions (read vertically):

```
1111111111111111111111111111111111111111111111111111
5555555555555555555555555555555555555555555555555555
4444455555555555555555555556666666666666666777777777777
9999900123334445688999900001112344566800111233466777
0456768413468239904567801234567659099674398906070060156
```

Haplotype	Cluster	Variable sites	Ast	Cco	Gar	Los	Mer	Pot	And	Car	Lus	Mar	Sor	NI	SI	Bar	Exm	Whole Dataset
H1	Ia	CCCGCGTAGATTACAACGCGAATGTTCGAATACGATGACTCCAAGTACCCTCT	4	8	2	1	1	1	8	2	5	4		17	19	6		42
H2	Ia	....T.........															1	1
H3	Ia	...........................A.........................					1					1		1	1			2
H4	Ia	...........................C.........................						1		1	1			1	2			3
H5	Ia	...........................CT..............C....T....						1						1				1
H6	Ia	...............A.....................................		2					1					2	1			3
H7	Ia	...............A.....A...............................	1		1		1		4					3	4	1		8
H8	Ia	.......A....C............C..								1					1			1
H9	Ia	......A....C.A............C..						1		1	1			1	2			3
H10	Ia	......A....CT................							1						1			1
H11	Ia	......G...A....A............C..							1						1			1
H12	Ia	.....A...									1				1			1
H13	Ib	...A...					1							1				1
H14	Ib	...AT..						1						1				1
H15	Ib	.T.A.............C...................................						1						1				1
H16	IIIa	.T.A......C......T...C..A.C..........................															7	7
H17	U	.T.A......C.......TA.....A.....C......G..............			1									1				1
H18	U	.T.A......C.......TA.....A.....A.G...................			1									1				1
H19	II	.T.A......C.......CT.....AG.A.................C..	1		1				2		5	2		2	9	1		12
H20	IIIb	.T.A......C.......C.T...A...........................					1							1			2	3
H21	U	.T.A......C........T.GG...A.....C........TCTC							2						2			2
H22	Va	.T.A......C......G...T...AG.............C..			1									1		1		2
H23	Va	.T.A......C......G...T...AG.....T.....................														2		2
H24	Va	.T.A......C......G...T...AG.....T......G.....														2		2
H25	U	.T.A......C......G...TA...A...G.C.....................														1		1
H26	IIIa	.T.A......C.......A....T...C..A.C....................															1	1
H27	IIIa	.T.A......C.......A....A.....C..C....................					1							1				1
H28	U	.T.A......C.......A....TA.....A.....C....A.G.....	2											2				2
H29	II	.T.A......C.......A....CT.....AG.A.............C..			1	2	1	1				1		5	1			6
H30	IIIb	.T.A......C.......A....C.T...A.....................	1	3	1							1		5	1			6
H31	IIIb	.T.A......C.......A....C.T...AG....................					2							2				2
H32	IIIb	.T.A......C.......A..AC.T...A.....................														1		1
H33	IIIa	.T.A......C.......A....T...C..A.C.....................															1	1
H34	II	.T.A......C.......TA.....CT....G.AG.A.............C..	1											1				1
H35	U	.T.A......C...GGTA...G...T...AG......C...G..........															1	1
H36	U	.T.A......C...T.....G...T...AG......................			1									1				1
H37	IV	.T.A......C..T......G...T.....TAG.A..C.........C..	1										4	1	4			5
H38	IV	.T.A......C..T......G...T.....TAG.A...C.T.......C..					2							2				2
H39	IV	.T.A......C..T......G...T.....TAG.A.TC........C..			1									1				1
H40	U	.T.A......C..T.....A.G...T.....AG.A..................										2	3		5			5
H41	IV	.T.A......C..T.....A.G...T.....TAG.A.............C..									3				3			3
H42	IV	.T.A......C..T.....A.G...T.....TAG.A..C...........C..	1		1					3				2	3			5
H43	IV	.T.A......C..T.....A.G.A..T.G...TAG.A..C.........C..															1	1
H44	Vb	.T.A......C.G.......G...T.....A.....T............		1										1				1
H45	Vb	.T.A......C.G....A......T.....AG.....T............					1		4	1				1	5	1		7
H46	Vb	.T.A......C.G..A.G...T.....AG.....T............					2							2				2
H47	U	.T.A.....-C........A....TA...A.............C....										1			1			1
H48	U	.T.A.....A.C........G......T...A............C...TC..														1		1
H49	U	.T.A.....A.C........A....T....A............C...TC..								1					1			1
H50	Ib	.TTA...							1						1			1
H51	Ib	.TTA....................C............................							1						1			1
H52	U	.TTA..............CT....A.............C..					2							2				2
H53	U	.TTA......C.......CT....A......G.....C..				1								1			3	4
H54	Ib	.TTA..C..............C...............................							1						1			1
H55	II	TT.A......C.......CT.....AG.A.............C..											1		1			1

			Ast	Cco	Gar	Los	Mer	Pot	And	Car	Lus	Mar	Sor	NI	SI	Bar	Exm	Whole Dataset
N			12	14	10	10	12	13	21	10	16	12	7	71	66	14	20	171
Nucleotide diversity			7.9	3.5	7.0	7.2	6.2	4.8	6.1	7.3	6.4	6.5	2.3	6.9	7.1	6.1	5.2	7.1
K			8	4	9	7	10	13	9	4	6	7	2	37	19	9	9	55
Haplotype diversity			66.7	28.6	90.0	70.0	83.3	100	42.9	40.0	37.5	58.3	28.6	52.1	28.8	64.3	45.0	32.4

Northern Iberian (NI) group includes Asturcón (Ast), Caballo de Corro (CCo), Garrano (Gar), Losino (Los), Mérens (Mer), and Pottoka (Pot) breeds; Southern Iberian (SI) group includes Andalusian (And), Carthusian (Car), Lusitano (Lus), Marismeño (Mar), and Sorraia (Sor) breeds. The non-Iberian breeds are Barb (Bar) and Exmoor (Exm).

Letter U in the cluster column means that the haplotype has not been assigned to any cluster identified in Figure 2.
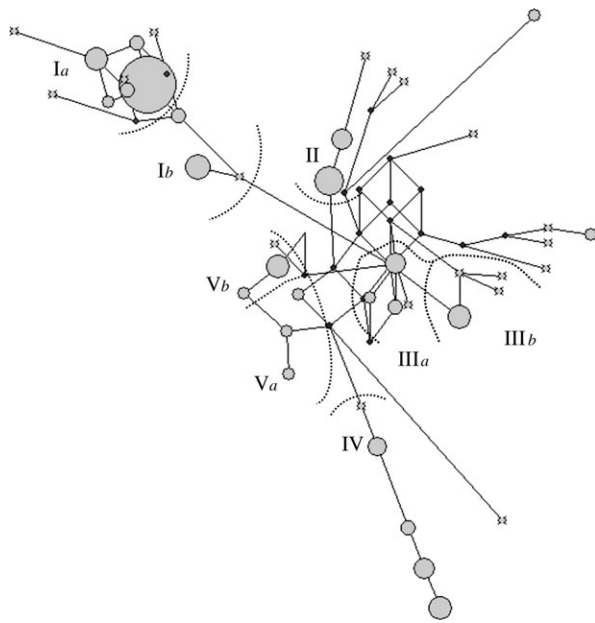
Figure 2. Median-joining tree constructed with maximum parsimony trees. The network has been contracted for character conflicts. The number of times in which each haplotype is present in the sample is proportional to the area of the circle. Small black circles indicate missing haplotypes. Sequence motifs defining major mtDNA clusters are as follows: CCG motif from np 15,494 to 15,496 for cluster Ia (whereas the others have TCA in the same positions); A15666 for cluster II; C15617 for cluster IIIa (includes most Exmoors); C15601 for cluster IIIb; T15542 for cluster IV (includes Sorraias); and T15709 for cluster V.

Asturcón and Caballo de Corro were the same population until recently (Álvarez Llana 1995; García-Dory 1980). Consequently, this can be explained by sampling due to possible genetic bottlenecks, random drift, or differential selection

pressure. Neighbor-joining trees were constructed based both on p distances and linearized Φ_{ST} values between breeds (not shown), but no differentiation patterns, geographical or phenotypical, could be assessed. Additionally, an AMOVA analysis was conducted using the Arlequin program (not shown). No statistical differentiation between Iberian groups was found, regardless of the inclusion of Barb and Exmoor breeds in the Southern Iberian and Northern Iberian groups, respectively.

Discussion

At a species level, horses possess high haplotype diversity in the D-loop region of mtDNA (Hill et al. 2002; Kim et al. 1999; Vilà et al. 2001). Our results show the existence of different genetic pools in Iberian horse breeds. These different mtDNA lineages are not preferentially distributed in some breeds within groups, leading to lack of differentiation between breeds and groups of breeds. The genetic diversity observed in Iberian horse breeds is generally unaffected by the recent genetic bottlenecks experienced by most sampled Iberian breeds. Recovery of the Northern Iberian breeds commenced in the 1980s from a few dozen individuals per breed (García-Dory 1980; Martínez Sanz et al. 1996). Carthusian individuals originated from fewer than 150 mares (Valera et al. 2005). We had expected to find a small number of mtDNA haplotypes in most Iberian breeds.

Results from the present analyses did not confirm the expected role of donor population for both Barb and Exmoor breeds. Moreover, Exmoor did not present consistent genetic relationships with Southern Iberian breeds. An overview of our results would confirm both the close genetic relationship between Iberian and Northern African horse populations and the slight genetic relationship between Exmoor and Northern Iberian breeds existing at the mtDNA level. Exmoor shared haplotypes with Losino and Mérens breeds. This may be the result of the well-documented use of British ponies on Losino and Pottoka populations during the first third of 20th century (Aparicio 1944). No

Table 2. Linearized Φ_{ST} distances (below diagonal) and p distance (above diagonal) between horse breeds

	And	Ast	Bar	Car	Cco	Exm	Gar	Los	Lus	Mar	Mer	Pot	Sor
And		0.0245	0.0202	0.0272	0.0164	0.0288	0.0239	0.0329	0.0229	0.0228	0.0224	0.0197	0.0356
Ast	0.0392		0.0237	0.0260	0.0217	0.0270	0.0235	0.0284	0.0240	0.0235	0.0234	0.0248	0.0277
Bar	0	0.0035		0.0257	0.0162	0.0257	0.0224	0.0303	0.0220	0.0218	0.0213	0.0202	0.0322
Car	0.2021	0.0083	0.1313		0.0244	0.0259	0.0248	0.0249	0.0251	0.0244	0.0239	0.0284	0.0234
Cco	0	0.1381	0	0.3655		0.0255	0.0212	0.0316	0.0203	0.0203	0.0190	0.0159	0.0346
Exm	0.4943	0.2346	0.3437	0.2389	0.6938		0.0234	0.0260	0.0264	0.0249	0.0221	0.0299	0.0253
Gar	0.0839	0	0.0146	0.0246	0.2222	0.1454		0.0264	0.0227	0.0217	0.0214	0.0244	0.0258
Los	0.4771	0.1112	0.3536	0.0116	0.8115	0.2552	0.1009		0.0266	0.026	0.0271	0.0338	0.0180
Lus	0.0800	0	0.0361	0.0803	0.2015	0.3428	0.0037	0.1574		0.0207	0.0225	0.0217	0.0263
Mar	0.0690	0	0.0193	0.0427	0.2050	0.2588	0	0.1219	0		0.0215	0.0223	0.0256
Mns	0.0697	0	0.0184	0.0427	0.1640	0.1435	0	0.1951	0.0504	0		0.0222	0.0288
Pot	0.0521	0.1562	0.0862	0.3907	0.1298	0.7383	0.2290	0.6805	0.1330	0.1607	0.1844		0.0381
Sor	1.1636	0.4967	1.0502	0.3744	2.3393	0.7836	0.5674	0.0874	0.6193	0.5919	0.8552	1.9002	

And = Andalusian, Ast = Asturcón, Bar = Barb, Car = Carthusian, Cco = Caballo de Corro, Exm = Exmoor, Gar = Garrano, Los = Losino, Lus = Lusitano, Mar = Marismeño, Mer = Mérens, Pot = Pottoka, Sor = Sorraia.

information on this topic is available for the Mérens breed; however, a similar use of British ponies on the breed cannot be excluded.

Furthermore, the hypothesis that the Sorraias represent the ancestors of the present Southern Iberian horses (d'Andrade 1945) would not be well supported on a maternal genetic basis. Haplotypes present in Sorraias do not cluster with the more frequent mtDNA haplotypes in Iberian horses (our cluster Ia), although they are not a genetic rarity in the total Iberian horse population. However, the hypothesis cannot be completely excluded; Sorraias have undergone a significant bottleneck, and only two mtDNA founder lineages remain in the breed (Luis et al. 2002), thus probably affecting our ability of ascertaining this task without uncertainty.

The present study highlights the wide diffusion on the Iberian Peninsula of the haplotypes of cluster Ia, which would roughly form the North African cluster. Two explanations are possible for the presence of this family of haplotypes in Iberian horses. First, we may think of this as the result of the recent introgression of Barb-type horses in the Iberian Peninsula during the Muslim invasions in the Middle Ages. The high frequency found for haplotype 1 (see Table 1) in Iberian samples would suggest a recent intensive introgression. Second, we could consider this haplogroup to be representative of the ancestral Iberian horse population.

Certain historical evidence allows us to reject the hypothesis of an introgression of Barb horses during the Muslim period after 700 CE as the sole explanation for the present situation of Iberian horses. The Muslim people who took part in the invasion of the Iberian Peninsula were small in number with respect to the existing human population (Chejne 1980; García de Cortázar 1988). The number of horses brought over by the Muslims has been estimated also to be small (Forbis 1976). Moreover, Ia haplotypes are present in moderate to high frequencies in Northern Iberian horse populations. Muslim control of the geographic area containing Northern Iberian horses was unstable and lasted only three decades (García de Cortázar 1988).

Despite the fact that the number of mtDNA haplotypes shared between Barb and Iberian horses would support the classical hypothesis suggesting that the ancestral Iberian horse population would basically be North African in origin (Aparicio 1944; Sotillo and Serrano 1985), archaeological findings and cave paintings suggest that light horses with convex head profiles (characteristic of present Barb and Southern Iberian horses) have been the most frequent in the entire Iberian territory (Aparicio 1944). Our data indicate the close genetic relationship, at a mtDNA level, between Northern African and Iberian horse populations. This geographical clusterization and extensive genetic diversity, together with the aforementioned archaeological findings, are consistent with the presence of wild mares for domestication in this geographical area (Jansen et al. 2002). However, our results do not clarify without uncertainty the size and direction of the gene flow between them. In this respect, the introgression of Iberian horses into North Africa has happened, at least, from Roman times to early 20th century

(Aparicio 1944), thus probably limiting our possibilities of clearly ascertaining gene flow patterns.

A major finding of our analysis is the lack of maternal genetic support for the traditional classification of the native Iberian populations of horses into Northern or Southern Iberia (Aparicio 1944; Sotillo and Serrano 1985). Phenotypic differentiation between Iberian horse groups suggests that Northern Iberian ponies are the result of a process of replacement rather than introgression (Alvarez Llana 1995; Aparicio 1944; Jordana et al. 1995; Sotillo and Serrano 1985). However, our results indicate the opposite pattern. Only a Mérens sample showed a haplotype included in our cluster IIIa. This cluster is consistent with the one classified as C1 by Jansen et al. (2002), which is characteristic of British and Northern European ponies. This scenario could support the hypothesis that the present Northern Iberian horses may be the result of a male-mediated introgression of Northern European ponies, probably linked with present British pony breeds. This is consistent with sex-biased ancient breeding practices (Vilà et al. 2001). A small number of male horses has probably been intensively used to form the present domestic horse (Lindgren et al. 2004). The extant mares have underwent intense selection (probably for type) thus leading to a male-mediated type differentiation that would not affect the maternal lineages (Vilà et al. 2001). This male-mediated selection could have been led by dominant human populations migrating from Central Europe to the Iberian Peninsula (Villar 1996). The impact on the mtDNA composition of the existing horse population could have been slight, depending on the size of the incoming group, particularly if it was predominantly male. A situation such as that described has been reported in African cattle in which zebu male-mediated interbreeding has produced major phenotypical changes in the extant taurine cattle without affecting mtDNA lineages (Bradley et al. 1996; Loftus et al. 1994). The hypothesis of the phenotypic selection of Northern Iberian horses by a male-mediated dominant human population predicts that the genetic relationships of these populations based on Y-chromosome markers should reflect the evident phenotypic relationships between Northern Iberian and British ponies more closely than that obtained from mtDNA. Extensive noncoding Y-chromosomal sequences have been screened for single nucleotide polymorphisms in a number of stallions, representing several domestic horse breeds (Lindgren et al. 2004; Wallner et al. 2003, 2004). Even though these studies did not identify a single polymorphic site, this will be an interesting topic for further research.

In conclusion, the present study reveals that the Iberian Peninsula is not a contact zone, at a mtDNA level, between Northern European and Northern African horse populations. Phenotypic differences among the Northern and Southern Iberian groups of breeds are not explained by population subdivision based on maternal lineages. The basis of this lack of genetic differentiation between Northern and Southern Iberian horse groups is that matrilines related to the Northern African cluster are the most common in Iberian horse breeds, regardless of geography or phenotype.

Acknowledgments

References

Álvarez Llana J, 1995. Morfología y caracteres raciales. In: Asturcones (Álvarez Llana J, Álvarez Sevilla A, and Jáuregui Campos J, eds). Oviedo, Spain: Ediciones Caja de Asturias; 3–39.

d'Andrade R, 1945. Sorraia. Boletim Pecuário 13:1–13.

d'Andrade R, 1954. Alrededor del caballo Español. Lisboa, Portugal.

Aparicio G, 1944. Zootecnia especial: etnología compendiada, 3th ed. Córdoba, Spain: Imprenta Moderna.

Bandelt H-J, Forster P, and Röhl A, 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Bandelt H-J, Forster P, Sykes BC, and Richards MB, 1995. Mitochondrial portraits of human populations. Genetics 141:743–753.

Bradley DG, MacHugh DE, Cunningham P, and Loftus RT, 1996. Mitochondrial diversity and the origins of African and European cattle. Proceeding of the National Academy of Science USA 93:5131–5135.

Chejne AG, 1980. Historia de España Musulmana. Madrid, Spain: Ediciones Cátedra.

Feng DF and Doolittle RF, 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. J Mol Evol 25:351–360.

Forbis J, 1976. The classic Arabian horse. New York: Liveright.

Forster P, Torroni A, Renfrew C, and Röhl A, 2001. Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. Mol Biol Evol 18:1864–1881.

García de Cortázar JA, 1988. Historia de España dirigida por Miguel Artola; 2, La época medieval. Madrid, Spain: Alianza Editorial.

García-Dory MA, 1980. Asturcón; Caballo de los astures. Oviedo, Spain: Ediciones Caja de Ahorros de Asturias.

Hill EW, Bradley DG, Al-Barody M, Ertugul O, Splan RK, Zakharov I, and Cunningham EP, 2002. History and integrity of thoroughbred dam lines revealed in equine mtDNA variation. Anim Genet 33:287–294.

Jansen T, Forster P, Levine MA, Oelke H, Hurles M, Renfrew C, Weber J, and Olek K, 2002. Mitochondrial DNA and the origins of the domestic horse. Proc Nat Acad Sci USA 99:10905–10910.

Jordana J and Parés PM, 1999. Relaciones genéticas entre razas ibéricas de caballos utilizando caracteres morfológicos. Prototipos raciales. AGRI 26:75–94.

Jordana J, Parés PM, and Sánchez A, 1995. Analysis of genetic relationships in horse breeds. J. Equ Vet Sci 7:320–328.

Kim K-I, Yang Y-H, Lee S-S, Park C, Ma R, Bouzat JL, and Levin HA, 1999. Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. Anim Genet 30:102–108.

Lindgren G, Backström N, Swinburne J, Hellborg L, Heinarsson A. Sandberg K, Cothran G, Vilà C, Binns M, and Ellegren H, 2004. Limited number of patrilines in horse domestication. Nature Genet 36:335–336.

Loftus RT, MacHugh DE, Bradley DG, Sharp PM, and Cunningham EP, 1994. Evidence for two independent domestications of cattle. Proceeding of the National Academy of Science USA 91:119–144.

Luis C, Bastos-Silveira C, Cothran EG, and Oom MM, 2002. Variation in the mitochondrial control region sequence between the two maternal lines of the Sorraia horse breed. Genet Mol Biol 25:309–311.

Martínez Sáiz JM, Valera M, and Molina A, 1996. El caballo Losino. AGRI 19:17–28.

Sambrook J, Fritsch EF, and Maniatis T, 1989. Molecular cloning: a laboratory manual, 2nd ed. New York: Cold Spring Harbor Laboratory Press.

Sbisà E, Tanzariello F, Reyes A, Pesole G, and Saccone C, 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205:125–40.

Schneider S, Roessli D, and Excoffier L, 2000. Arlequin v.2.0: a software for population genetics data analysis. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.

Slatkin M, 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462.

Sotillo JL and Serrano V, 1985. Producción animal I. Etnología zootécnica. Madrid, Spain: Ediciones Tebar-Flores.

Valera M, Molina A, Gutiérrez JP, Gómez J, and Goyache F, 2005. Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. Livest Prod Sci 95:57–66.

Vilà C, Leonard JA, Götherstöom A, Marklund S, Sandberg K, Lindén K, Wayne RK, and Ellegren H, 2001. Widespread origin of domestic horse lineages. Science 291:474–477.

Villar F, 1996. Los Indoeuropeos y los orígenes de Europa: lenguaje e historia. Madrid, Spain: Editorial Gredos.

Wallner B, Brem G, Müller M, and Achmann R, 2003. Fixed nucleotide differences on the Y chromosome indicate clear divergence between *Equus przewalskii* and *Equus caballus*. Anim Genet 34:453–456.

Wallner B, Piumi F, Brem G, Müller M, and Achmann R, 2004. Isolation of Y chromosome-specific microsatellites in the horse and cross-species amplification in the genus *Equus*. J Hered 95:158–164.

Xu X and Arnason U, 1994. The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. Gene 148:357–362.