



ORIGINAL ARTICLE

## Effects of a 10-year conservation programme on the genetic diversity of the Pottoka pony – new clues regarding their origin

F. Rendo, M. Iriondo, C. Manzano & A. Estonba

Department of Genetics, Physical Anthropology and Animal Physiology; Faculty of Science and Technology, University of the Basque Country, Bilbao, Spain.

### Keywords

conservation programme; European pony; genetic diversity; Horse; microsatellite; Pottoka breed.

### Correspondence

Andone Estonba, Department of Genetics, Physical Anthropology and Animal Physiology; Faculty of Science and Technology, University of the Basque Country, E-48940 Bilbao, Spain.  
Tel.: +34 94 601 55 17; Fax: +34 94 601 31 45; E-mail: andone.estonba@ehu.es

Received: 4 October 2010;

accepted: 1 August 2011

### Summary

Here, we present the results of a genetic analysis of 463 Pottoka ponies corresponding to four generations, using 17 microsatellite markers. Ten years after the beginning of the Pottoka conservation programme, the values for the genetic diversity of the breed are still high and stable, indicating the success of the programme. We found null alleles in Pottoka for the ASB23, HMS3 and HTG10 microsatellites. Together with information obtained from other pony breeds from the Iberian Peninsula, this finding indicates that these microsatellites should not be used for phylogenetic analyses or parentage tests, at least for these breeds. The high heterozygosity exhibited by this breed in comparison to other ponies, together with its genetic proximity to the centroid of the allele frequencies, suggest that Pottoka allele frequencies are close to those initially exhibited by the ancestors of current European ponies. The results obtained in the current work, together with results from previous studies of ponies and horses from the Iberian Peninsula, corroborate the idea of a unique origin of all ponies from the European Atlantic Area. In contrast, our results do not corroborate the idea that these are derived from a domestication event in the Iberian Peninsula, nor that they have incorporated ancient Iberian horse genes into their genetic pool to a larger extent than other horse breeds.

### Introduction

Pottoka is a semi-feral pony breed which inhabits the Basque Country (Western Pyrenees). It was catalogued by the Food and Agricultural Organization (FAO) as an endangered breed, since in 1995, there were only 400 females and 170 males. Moreover, Pottoka presented a decreasing population trend. The establishment of the Pottoka studbook in 1995 contributed to reversing this trend, with the result that in 2009, there were around 685 animals (610 mares and 75 studs) which were classified as being purebred (data from the Basque Federation of Breed-

ers of Pottoka Basque Ponies – EPOFE, personal communication). According to the breed standard, the Pottoka is a pony with a height of 1.15–1.32 m. at the withers, with a black or dark-bay coloured coat, a head with a sub-concave profile, small ears and large nasal orifices, a thick and strong neck with long mane, sub-concave back and with a thick coat in Winter (EPOFE, <http://www.pottoka.info/index.php?id=en>; Figure 1).

These animals freely reproduce within their herd, but EPOFE certifies, within each herd, the appropriateness of the males as well as the females which are to be used for reproductive ends. Each breeder can



**Figure 1** A typical purebred male Pottoka pony.

have its own male, or EPOFE can provide its own males to breeders who request them. On the other hand, there is a programme to restock males for their subsequent selection and use as reproducers. In the first place, at 3 years of age, EPOFE morphologically grades all the future studs and carries out a paternity test to verify their genealogy. Those animals which present a higher morphological rating and whose genealogy is corroborated by the genetic paternity tests are selected as future candidate studs. At approximately 5 years of age, animals with the highest grading based on morphological evaluation are selected as the studs offered by EPOFE to breeders who request them. Thus, the principal differences with respect to breeding prior to the publication of the studbook in 1995 consist of herds being formed exclusively by purebred individuals and of the use of certified reproductive males in those herds which require them.

The domestication of the horse, which began over 6000 years ago (Outram *et al.* 2009), seems to have been a complex process. In present-day horses, high matrilineal diversity has been reported (Vilà *et al.* 2001; Jansen *et al.* 2002; Cieslak *et al.* 2010), which has been attributed to their multiple origins, an extremely large number of female founders and/or a large-scale introgression of local lineages into the domestic stock (Cieslak *et al.* 2010). In contrast, very low levels of Y-chromosome variability have been found (Lindgren *et al.* 2004). Additionally, the extensive mobility of the horse, favoured by its domestic roles, is also responsible for the progressive obscuring of the genetic structure of the species during the migrations subsequent to its domestication (Hill *et al.* 2002). Recently, a high quality draft of the sequence of the equine genome has been published,

revealing a pattern of linkage disequilibrium which is compatible with an early domestication of widespread herds of females and an intimate relation among all equine breeds worldwide (Wade *et al.* 2009). Regarding the origin of Iberian breeds, two studies of the origin and history of mitochondrial DNA (mtDNA) lineages (Cieslak *et al.* 2010; Lira *et al.* 2010) have raised the possibility of the existence of an independent domestication episode in the Iberian Peninsula, or at least, the use of wild Iberian maternal lineages in a process of restocking populations. In this regard, a study employing nuclear markers (Warmuth *et al.* 2011) has suggested that the Iberian Peninsula as well as the steppes of Eastern Europe may have been refuges during the Holocene, from which were derived the equine populations of the central and northern Europe. It cannot be concluded from these results that the Iberian Peninsula was an area of independent domestication; nevertheless, in keeping with the mtDNA results, these findings do indicate that the genetic contribution of these wild populations to local domestic stocks may have been considerable.

Regarding genetic studies of the Pottoka breed, Cañón *et al.* (2000) showed using microsatellite markers that this breed grouped with the other pony breeds from the northern Iberian Peninsula, although with a confusing population structure. These authors suggested that this was owing to a common origin of the analysed pony breeds which, together with a similar morphology, would have permitted in the past a reciprocal introgression between these breeds. On the other hand, the results obtained by Royo *et al.* (2005) upon analysing the mtDNA of Iberian Peninsula breeds, did not manage to differentiate Pottoka from the other pony breeds; neither did they support the traditional classification of native Iberian breeds into northern and southern breeds. Finally, Solís *et al.* (2005) analysed the Pottoka breed within the context of other autochthonous breeds in the western Pyrenees region (such as the Jaca Navarra pony, and the 'Euskal Herriko Mendiko Zaldia' and Burguete heavy horses). The Pottoka breed was found to be situated at one of the ends of the spectrum of regional genetic variability, and the Jaca Navarra was found to be in a position between Pottoka and the heavy horses. The within breed genetic variability detected for Pottoka and Jaca Navarra was found to be very similar to that found for these same breeds by Cañón *et al.* (2000), being slightly higher than that reported for other European or American horse breeds (Solís *et al.*, 2005).

In the light of the above, the objective of the current study was double: on the one hand, and follow-

ing more than 10 years of implantation of the Pottoka conservation programme, we aimed to characterize the current genetic status of the breed, thereby evaluating the management to which it is being subjected. On the other hand, we aimed to enhance our understanding of the origin of this breed, by providing new data regarding its genetic status within the context of other breeds of Iberian ponies in particular, and of European ponies in general.

## Materials and methods

### Samples

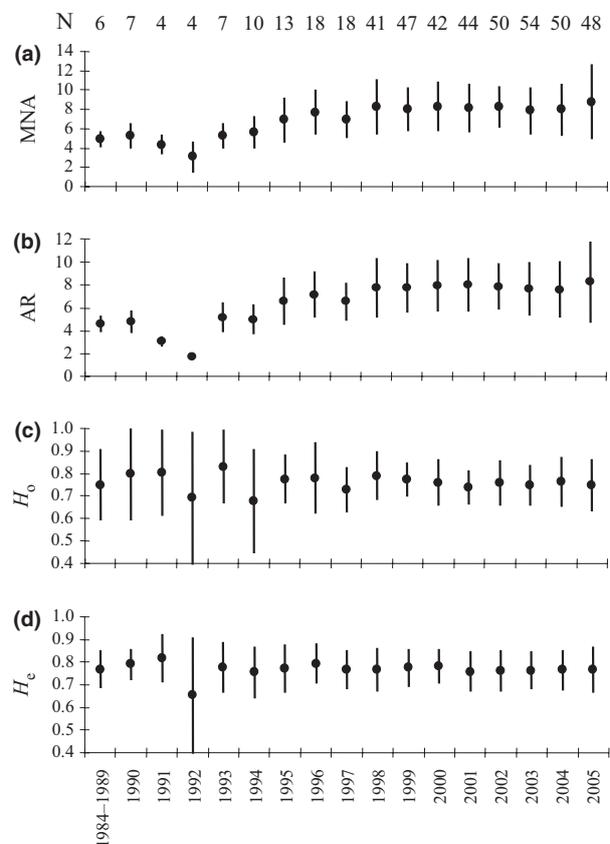
Blood samples were taken from a total of 463 purebred Pottoka ponies, from the Gipuzkoa province, born between 1984 and 2005. Sampling of the horses was generally carried out at 3 years of age, when morphological and genetic grading of the individuals is performed. Seventeen age subgroups were established: one per year from 1990 to 2005, both inclusive, with the exception of those born from 1984 to 1989, which were grouped together owing to the low number of available individuals. The sample size of each subgroup is indicated in Figure 2. The 376 individuals born after 1998 represent 76.9% of the 489 animals currently recorded in the Gipuzkoa studbook.

### Genetic analysis

Genomic DNA was extracted from whole blood using the DNeasy 96 Blood & Tissue kit (Qiagen, Germantown, MD, USA) and an automatic STARlet (Hamilton, Bonaduz GR, Switzerland) robot. We analysed the microsatellite markers pertaining to the 'Stock-Marks for Horses' Genotyping Kit panel (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA). These include the following markers: AHT4, AHT5, ASB2, ASB17, ASB23, CA425, HMS1, HMS2, HMS3, HMS6, HMS7, HTG4, HTG6, HTG7, HTG10, LEX3 and VHL20. All microsatellites were simultaneously detected using ABI PRISM 3100 Avant and 3130 XL (Applied Biosystems) genetic analyzers. The GeneScan<sup>®</sup> 3.7.1, Genotyper<sup>®</sup> 3.7 and GeneMapper programs (Applied Biosystems) were used for the analysis and genotyping of microsatellites. Individual genotypes were used to certify the genealogy of all the individuals inscribed in the Pottoka studbook.

### Statistical analysis

The mean number of alleles (MNA), allelic richness (AR), observed ( $H_o$ ) and expected ( $H_e$ ) het-



**Figure 2** Sample size of each age subgroup and variation in mean number of alleles (MNA), allelic richness (AR), observed heterozygosity ( $H_o$ ) and expected heterozygosity ( $H_e$ ) throughout the studied generations.

erozygosity, as well as the  $F_{IS}$  and  $F_{ST}$  parameters were calculated with the FSTAT program (Goudet 1995). The combined non-exclusion probability of identity ( $PNE_{ID}$ ) and of the parent pair ( $PNE_{PP}$ ) were calculated using the CERVUS 3.0.3 program (Kalinowski *et al.* 2007). Confidence intervals were determined by jackknifing, and statistical significance was considered after 15 000 permutations. Hardy-Weinberg equilibrium (HWE) was tested by means of the Exact Test option implemented in the GENEPOP 4.0 software (Rousset 2008). In cases in which it was deemed necessary, the Bonferroni correction was introduced. In particular, the probability values  $p < 0.0030$  and  $p < 0.0037$  were considered significant for 17 microsatellite *loci* or 17 population subgroups and for 14 microsatellite *loci*, respectively. Using the MICRO-CHECKER v2.2.3 program (Van Oosterhout *et al.* 2004), we tested the potential presence of non-amplified alleles (null alleles).

To evaluate the relative quantity of gene flow experienced by each breed, we applied the method proposed by Harpending & Ward (1982), based on the standardized matrix of variance–covariance of allele frequencies. To this end, we obtained the allele frequencies of seven microsatellites (AHT4, AHT5, HMS6, HMS7, HTG4, HTG6 and VHL20) from another 14 pony breeds: Garrano (Luis *et al.* 2007a) from Portugal, Jaca Navarra (Solis *et al.* 2005), Asturcón, Cabalo Galego and Losino (Checa 2004) from Spain, Landais and Poney Français de Selle from France (Leroy *et al.* 2009), and New Forest, Welsh (Leroy *et al.* 2009), Dales, Fell, Shetland, Exmoor and Connemara (Luis *et al.* 2007a) from the UK and Ireland. According to this method, a simple linear relation can be expected between the heterozygosity of each population ( $h_i$ ) and the distance of each population from the centroid of allele frequency ( $r_i$ );  $h_i = H(1 - r_i)$ , with  $H$  being the heterozygosity of the totality of all the populations. Expected heterozygosity ( $h_i$ ) and its standard deviation were calculated from allele frequencies according to Nei (1987). Distance from the centroid was calculated as  $r_i = (p_{ik} - P)^2 / P(1 - P)$ , where  $p_{ik}$  is the frequency of the allele  $k$  in the breed  $i$ , and  $P$  is the frequency of the allele  $k$  in the overall population of all breeds together.

To measure concordance between the matrices of geographical and genetic distances between breeds, first order correlations were calculated. Geographic distances were calculated considering the minimal distance by road between the places of origin of the breeds, while the  $F_{ST}$  genetic distances were taken from Reynolds *et al.* (1983), calculated using PHYLIP 3.5 software (Felsenstein 1989). Matrix comparisons were performed by the Mantel method, using the ZT program (Bonnet & Van de Peer 2002), obtaining significant values after 1 000 000 iterations.

## Results

The MNA values of the 17 microsatellites varied from 4.18 for HTG7 to 11.82 for ASB17 (Table 1). All microsatellites presented high heterozygosity. Minimal values of observed and expected heterozygosity were observed for the HMS1 microsatellite, whereas highest values were associated with ASB17. All microsatellites were in HWE, with the exception of ASB23, HMS3 and HTG10, which present an excess of homozygotes and a positive and significant  $F_{IS}$  after Bonferroni correction (Table 1). Analysis of these three microsatellites with the MICRO-CHECKER program revealed null alleles for all of them. More-

over, the paternity tests carried out for the male restocking programme corroborated the presence of null alleles: in the 165 cases in which biological compatibility (with 0 or 1 incompatibilities) was determined, we detected 26 cases in which the offspring was apparently homozygous, with one of the progenitors apparently being homozygous for another allele. All the single incompatibilities presented by the said progenitors corresponded to one of the three microsatellites (ASB23, 14 cases; HMS3, nine cases; and HTG10, three cases). These cases are indicative of the existence of null alleles, with the offspring and incompatible progenitor being heterozygous for the null allele. Consequently, these three microsatellites were not considered in subsequent analyses. For the 14 remaining microsatellites, we obtained a  $PNE_{ID}$  value of  $1.75E-16$  and a  $PNE_{PP}$  value of  $4.51E-10$ .

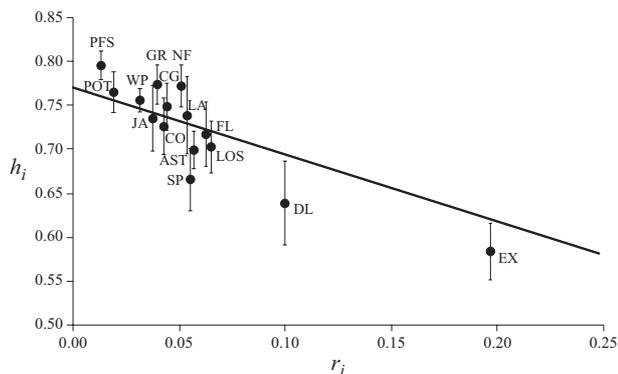
The temporal evolution of MNA, AR,  $H_o$ ,  $H_e$  and  $F_{IS}$  for the 14 selected microsatellites is represented in Figure 2. Because samples with more than 40 individuals are available only since 1998, data from previous years should be interpreted with caution. The homogeneity of values throughout the analysed years is particularly evident. Values are always very close to the mean (mean  $MNA_{1998-2005} = 8.20 \pm 0.28$ ; mean  $AR_{1998-2005} = 7.86 \pm 0.21$ ; mean  $H_o_{1998-2005} = 0.760 \pm 0.016$ ; mean  $H_e_{1998-2005} = 0.767 \pm 0.007$ ; mean  $F_{IS_{1998-2005}} = 0.009 \pm 0.019$ , N.S.). The  $F_{ST}$  between the 17 age groups was not found to be significant ( $F_{ST} = 0.002 \pm 0.001$ , N.S.). According to these microsatellite analyses, the values of genetic diversity for the Pottoka breed are situated at the higher end of the range observed for horse breeds (Cañón *et al.* 2000; Solis *et al.* 2005; Luis *et al.* 2007a; Leroy *et al.* 2009).

With a view to genetically comparing the Pottoka breed with other horse breeds, we took into account only those individuals born between 2003 and 2005 ( $N = 153$ ) to eliminate effects owing to generational overlapping, because females usually begin to reproduce from 3 years of age onwards. We analysed the relative amount of gene flow experienced by each breed, and found that the Exmoor, Dales, Shetland and Asturcón breeds were significantly below the expected heterozygosity line, i.e. their heterozygosity is less than that expected for the degree of differentiation which they present with respect to the centroid of the allele frequencies. The opposite occurs in the case of the New Forest, Garrano and Poney Français de Selle breeds, which were significantly above the line (Figure 3). Of all the breeds situated on the prediction line, Pottoka presents the lowest  $r_i$

**Table 1** Mean number of alleles (MNA), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), allelic richness (AR), Fisher exact test for Hardy–Weinberg equilibrium HWE and  $F_{IS}$  values and their significance for analysed populations

Microsatellite	MNA	$H_o$	$H_e$	AR	HWE	$F_{IS} \pm SE$
AHT4	7.00	0.779	0.802	6.55	N.S.	0.018 $\pm$ 0.017
AHT5	6.94	0.807	0.817	6.38	N.S.	0.018 $\pm$ 0.026
ASB2	7.94	0.802	0.801	7.37	N.S.	0.003 $\pm$ 0.025
ASB17	11.82	0.900	0.898	10.6	N.S.	0.011 $\pm$ 0.011
ASB23	7.00	0.662	0.842	6.60	*	0.246 $\pm$ 0.029*
CA425	6.53	0.719	0.764	6.14	N.S.	0.039 $\pm$ 0.019
HMS1	4.94	0.595	0.595	4.69	N.S.	-0.040 $\pm$ 0.040
HMS2	6.35	0.841	0.810	6.14	N.S.	-0.022 $\pm$ 0.023
HMS3	6.82	0.696	0.801	6.53	*	0.167 $\pm$ 0.037*
HMS6	5.94	0.743	0.781	5.67	N.S.	0.061 $\pm$ 0.028
HMS7	5.71	0.747	0.731	5.35	N.S.	0.006 $\pm$ 0.029
HTG4	5.82	0.702	0.709	5.40	N.S.	0.001 $\pm$ 0.029
HTG6	6.06	0.670	0.662	5.45	N.S.	-0.004 $\pm$ 0.038
HTG7	4.18	0.676	0.665	3.87	N.S.	-0.007 $\pm$ 0.024
HTG10	7.88	0.721	0.811	7.36	*	0.184 $\pm$ 0.027*
LEX3	7.82	0.803	0.836	7.58	N.S.	0.052 $\pm$ 0.026
VHL20	8.24	0.845	0.859	7.58	N.S.	0.003 $\pm$ 0.023

\*Significant after Bonferroni correction.



**Figure 3** The figure represents heterozygosity ( $h_i$ ) of the pony breeds with respect to distance from the centroid of the allele frequencies ( $r_i$ ) (Harpending & Ward 1982). The diagonal line corresponds to  $h_i = H(1 - r_i)$ , where  $H$  is the Nei heterozygosity (1987) for the total population. Analysed breeds are: Asturcón (AST), Cabalo Galego (CG), Connemara (CO), Dales Pony (DL), Exmoor (EX), Fell Pony (FL), Garrano (GR), Jaca Navarra (JA), Landais (LA), Losino (LOS), New Forest (NF), Poney Français de Selle (PFS), Pottoka (POT), Shetland Pony (SP) and Welsh Pony (WP).

value, while Dales and especially Exmoor exhibit the highest  $r_i$  values, i.e. a very large degree of differentiation with respect to the rest of the breeds. A significant correlation ( $r = 0.321$ ;  $p = 0.009$ ) was found upon comparing genetic and geographical distances (two by two) for all the pony breeds studied in this work, excluding those two breeds which presented highest genetic differentiation (Exmoor and Dales).

This correlation was still significant when only Exmoor was excluded ( $r = 0.245$ ;  $p = 0.018$ ).

## Discussion

Results obtained using 17 microsatellites, both for  $F_{IS}$  and for the analysis of genotypes, are indicative of the presence of null alleles in the ASB23, HMS3 and HTG10 microsatellites. The paternity tests which we carried out corroborate this finding, because in 26 cases, we observed the appearance of a single exclusion for one of the three microsatellites, with the foal and presumed progenitor being homozygous for different alleles. These cases are indicative of the existence of null alleles, with the offspring and progenitor who exhibits the incompatibility being heterozygous for the null allele. The existence of null alleles may also explain the fact that the HMS3 microsatellite is the only one to present a positive and highly significant  $F_{IS}$  value in the Asturcón breed (Checa *et al.* 1998), as well as in five breeds from the Cantabria region and two breeds from the Balearic Islands (Cañón *et al.* 2000). This may of course also be the case for other breeds, in addition to ponies in the Iberian Peninsula. This is indeed the case of the Lipizzan horse in which the presence of null alleles in microsatellites HMS3 and ASB2 was reported, owing to base substitutions in the sequences which flank the microsatellites, which are targets for the primers necessary for amplification (Achmann *et al.* 2001). The excess in homozygotes detected in two

horse breeds in Poland, for HTG10 (Ząbek *et al.* 2005), may also reflect the presence of silent alleles. Thus, the presence of null alleles would not only be related to particular microsatellites, but also to the analysed horse population, in such a manner that each breed could present null alleles in zero, one or various microsatellites. This result is of particular importance when comparing breeds with a view to determining the phylogenetic relationships among them, because the presence of null alleles in some of the breeds could generate misleading information regarding allele frequencies. The findings are also relevant for genealogical controls, because false exclusions can be avoided in the light of the effect of silent alleles on genotypes. In this study, once we detected the presence of null alleles for ASB23, HMS3 and HTG10 in Pottoka, these microsatellites were excluded from subsequent analyses.

One of the objectives of the current work was to genetically evaluate the impact on the Pottoka breed of specific breeding management practices, which have been in operation via a conservation programme since 1995. This breed is distributed among the Basque provinces of Gipuzkoa, Araba and Bizkaia. Although we have only analysed animals from the province of Gipuzkoa in the current study, it is reasonable to extrapolate the results to the entire breed for several reasons. First, all Pottoka animals are registered in the same studbook and are managed uniformly by the same organization (EP-OFE). Second, of all the 685 purebred animals which have been censused, 489 (71.4%) are in the Gipuzkoa province; the practical totality of all these animals in Gipuzkoa have been analysed in the current study. And third, Solís *et al.* (2005) did not detect genetic differences between the Pottoka purebred populations from Gipuzkoa and Bizkaia provinces. Results obtained with individuals who were born before 1995 should be interpreted with caution, as such cases were not very numerous. Results obtained with animals born since 1995 indicate that MNA, AR,  $H_c$  and  $H_o$  values remain around the mean (Figure 2). The slight increase in the MNA value after the first 4 years of existence of the Pottoka studbook (1995–1998) is likely owing to the increase in sampling size, as the MNA value is directly influenced by the number of sampled individuals. Mean values of observed and expected heterozygosity were not substantially different (Figure 2), which is consistent with the low  $F_{IS}$  values obtained ( $F_{IS} 1998-2005 = 0.009 \pm 0.019$ ; N.S.;  $F_{IS} 1984-2005 = 0.013 \pm 0.039$ ; N.S.). These results indicate the absence of both appreciable consanguin-

ity and within population structuring in the Pottoka breed from the Gipuzkoa province. Moreover, when serum proteins were analysed, an excess of homozygotes was not detected in the Bizkaia province (Arruga *et al.* 2001). Overall, these results indicate that the management to which the Pottoka breed has been subjected is correct. Thus, the qualification of offspring at 3 years of age, paternity tests, the maintenance of herds formed exclusively by purebred individuals, together with the programme of restocking of reproductive males for use in directed mating, generation after generation, have avoided the loss of genetic diversity which could have endangered the future of the breed. In this regard, it is noteworthy that the process of enhancement of the Pottoka breed includes the elimination of individuals with a chestnut coat, as these are not included in the Pottoka studbook. This exclusion of chestnut individuals from reproduction since 1995 does not seem to have affected any of the calculated parameters of diversity. A recent study (Rendo *et al.* 2009) has suggested that even the exclusion from reproduction of individuals who are carriers of the recessive alleles responsible for a chestnut coat colour, would not affect the expected heterozygosity of this breed.

Regarding phylogenetic relations between horse breeds, a number of factors hinder an unequivocal determination of the origin and parentage of the distinct horse breeds. These include the relatively recent origin of the domestic horse (Outram *et al.* 2009), considerable genetic contributions of wild horses to domestic horses (Cieslak *et al.* 2010; Warmuth *et al.* 2011), mixing of different breeds, and processes of genetic drift (Solís 2005). Thus, some breeds have been created from very few individuals, such as the English Thoroughbred or the Arabian breed. Others have experienced important 'bottleneck' and founder effects, such as those which occurred in risk breeds, recovered from very few individuals. This is the case of the Retuertas (Vega-Pla *et al.* 2006), Sorraia (Luis *et al.* 2007b) or Przewalski horses (Boyd & Houpt 1994). Ponies, in contrast, are breeds which have not been subject to intense selection, but it appears that they may have suffered the consequences of 'bottleneck' effects owing to the reduction of their effective population sizes. This reduction would have been owing to factors such as the loss of their habitats since the XVIII century, the mechanization of agriculture and the declining of shepherding during the XX century (Solís *et al.* 2005). Ponies may also have been crossed with other horse breeds with the objective of

increasing their size and of conferring on them a higher economic value. Nevertheless, if this mix had been substantial, they would of course no longer be ponies. This is the case of the *Euskal Herriko Mendiko Zaldia* and Burguete breeds, which can be assumed to have arisen originally from Pottoka and Jaca Navarra breeds respectively, following a crossbreed with males of heavy French breeds, such as Bretón, Percherón, Ardennais and Comtois (Solís 2005). For these reasons, the joint genetic analysis of a large number of breeds with different population histories, has resulted, in many cases, in the absence of distinctive genetic patterns which could explain parentage among them (Cañón *et al.* 2000; Vilà *et al.* 2001; Jansen *et al.* 2002; Outram *et al.* 2009). A similar situation was encountered when the phylogenetic relations between bovine breeds with different degrees of artificial selection were analysed (Rendo *et al.* 2004). For all these reasons, in this work in which we aimed at deepening our knowledge of the origin of the Pottoka breed, we included in our analyses only European pony breeds, with a view to facilitating the detection of the underlying genetic structure.

The elevated expected heterozygosity values obtained for the Pottoka breed, and the generalized presence of elevated values in the pony breeds from the Euro-Atlantic Area (northern Iberian Peninsula, France and the British Isles), is close to what would be expected for breeds which have been subjected to less selection and which are thus closer to wild ancestry (Lira *et al.* 2010; Warmuth *et al.* 2011). Regarding the analysis of the relative quantity of gene flow experienced by each breed (Harpending & Ward 1982), four breeds (Asturcón, Shetland, Dales and Exmoor) are located below the line of theoretical prediction in Figure 3. The heterozygosities of these populations are significantly lower than those expected in terms of the degree of differentiation which they present, relative to the centroid of allele frequencies. This fact would indicate that these breeds have experienced substantial genetic isolation, with considerable genetic drift and/or 'bottleneck' effects. In fact, the Exmoor breed had been restricted to 46 mares in 1940 (Jansen *et al.* 2002), while the size of the Asturcón population was dramatically reduced after the Spanish Civil War; this breed had to be recovered towards 1970 from only 21 reproductive black coat animals (Royo *et al.* 2007). The Dales breed was also on the verge of extinction, as it had been frequently used in the two world wars and its recuperation began around 1964 (Dales Pony Society, <http://www.dalespony.org>). The Shetland breed could have been

affected by a founder effect, genetic drift and/or a 'bottleneck' effect; these situations are observed in geographically isolated populations, such as those which are found on islands (Plante *et al.* 2007). In contrast, the Poney Français de Selle, Garrano and New Forest breeds present a heterozygosity which is significantly higher than expected, indicating that gene flow from other breeds not included in the current work has been considerable throughout its history (Harpending & Ward 1982). In this regard, the Poney Français de Selle breed arose from French pony females crossed with Arabian Thoroughbred, English Thoroughbred, Connemara, New Forest and Welsh studs (Association Nationale du Poney Français de Selle et du Poney de Sport – A.N.P.F.S.; <http://www.anpfs.com/>). There is evidence that a number of New Forest breeders tried to improve the aptitude of this breed by introducing Welsh, English Thoroughbred, Arabian Thoroughbred and Hackney breeds (The New Forest Pony Breeding & Cattle Society; <http://www.newforestpony.com/>). Regarding the Garrano breed, Portas *et al.* (2001) indicate that since 1940, the population has been reduced, and at the same time modified, by the introduction of exotic breeds to improve its size for meat consumption. The remaining breeds are situated near the line of theoretical prediction and it is noteworthy that of all these remaining breeds, Pottoka presents the lowest value of difference from the centroid.

These results taken together are compatible with the idea that all the analysed pony breeds have a common origin and that migration from other breeds does not seem to have exerted any appreciable effect on them. Thus, genetic drift would have been the principal evolutionary mechanism in the moulding of the allele frequencies which we currently observe in these breeds. This being the case, the Pottoka breed would be the least affected by genetic drift, i.e. it would be the breed which maintained the largest effective population size throughout its history. If the effect of genetic drift is analysed in a sufficiently large number of populations, no change in the mean of allele frequencies is expected. Thus, Pottoka allele frequencies would be most similar to those exhibited by the original pony population, before divergence into the present breeds occurred. The existence of a significant correlation between genetic and geographical distances is consistent with a single origin of all the analysed pony breeds, as has been suggested by authors such as Cañón *et al.* (2000) and Leroy *et al.* (2009). Moreover, this correlation suggests that isolation by geographical distance is the

principal mechanism which has moulded the allele frequencies of Euro-Atlantic Area pony breeds.

In general, in the trees constructed from genetic distances, ponies are situated in a central position with respect to other breeds of domestic horses who have been subjected to more intense selection (Leroy *et al.* 2009). These results, together with the distribution of mtDNA haplotypes (Vilà *et al.* 2001; Royo *et al.* 2005) indicate that ponies have not arisen from a distinctive domestication process, in contrast to other horse breeds. These same studies suggest that during the expansion of the domestic horse and the incorporation of local mares *en route* (Jansen *et al.* 2002), ponies did not experience a significantly larger wild contribution than other breeds, because if this were the case, ponies would have to present a considerable degree of differentiation from the other breeds; this is clearly not the case.

Regarding breeds from the Iberian Peninsula, Lira *et al.* (2010) and Cieslak *et al.* (2010) proposed that wild local mares were used during the initial processes of domestication in this area, or at least, they were used for processes of population restocking. This conclusion was arrived at principally on the basis of the identification of group C haplotypes (typical of the Lusitano horse) in Bronze Age samples from the Atapuerca archaeological site (northern Iberian Peninsula). These are a group of haplotypes which to date have only been found in the Iberian Peninsula. In addition, Warmuth *et al.* (2011), investigating patterns of genetic diversity in 24 European horse breeds by means of 12 microsatellite *loci*, showed that this genetic contribution of Iberian wild stock to local domestic horses might have been substantial. However, the C haplogroup (following the nomenclature used in Vilà *et al.* 2001 and Jansen *et al.* 2002) was not detected in local horses in the Basque Country and Navarra (Solís 2005). Thus, it seems that the process of an extensive use of local Iberian wild horses in establishing and/or restocking local domestic populations may not be applicable to the Pottoka breed, despite the fact that in the region which this horse currently inhabits, the horse would probably have been present in an uninterrupted manner for over a 100 000 years (Altuna & Mariezkurrena 2009). The absence of the Lusitano group C haplotypes in Pottoka (Royo *et al.* 2005; Solís 2005), together with the results of the current study, which indicate that this breed had a genetic contribution which was similar to that of the other pony breeds, suggest that the Palaeolithic horses of the region did not contribute to a significant extent to the Pottoka genetic pool.

Thus, our results point more to a unique origin of all the ponies of the Euro-Atlantic Area; they suggest that these did not arise from a domestication event in the Iberian Peninsula, and that they have not incorporated ancient Iberian horse genetic material into their genetic pool to a larger extent than other horse breeds. On the other hand, the detection in the only sample from the Middle Ages, from the Atapuerca site, of a haplotype which was absent in the Bronze Age sample (Lira *et al.* 2010), leads these authors to propose the entrance of horses into the Iberian Peninsula in historical times. In this context, the origin of ponies from the Northern Iberian Peninsula, including Pottoka, could be the result of the introduction of domestic horses into the Iberian Peninsula, at some stage between the Bronze Age and the Middle Ages.

### Acknowledgements

The authors would like to express their thanks to the staff of the research team, to Ainhoa Solís for her invaluable prior analyses, and to Irati Miguel, from the Sequencing and Genotyping Unit of the Genomic Service – Bizkaia (SGIKer-UPV/EHU) for her technical assistance. We would also like to thank the collaboration of INRA and LABOGENA who provided us with the allele frequencies of some of the analysed breeds. We are indebted to the Basque Federation of Breeders of Pottoka Basque Ponies/Euskadiko Pottokazaleen Federazioa (EPOFE), and in particular to Narciso Arrillaga, Ainhoa Egia, Jose Javier Ormazabal and Mirari Inza, for their collaboration in obtaining samples and for the information provided. This work has been partially financed by the Department of Agriculture and Fisheries of the Basque Government (PA 08/05).

### References

- Achmann R., Huber T., Wallner B., Dovc P., Müller M., Brem G. (2001) Base substitutions in the sequences flanking microsatellite markers HMS3 and ASB2 interfere with parentage testing in the Lipizzan horse. *Anim. Genet.* **32**, 52.
- Altuna J., Mariezkurrena K. (2009). El caballo al final de la última glaciación en el período postglacial del País Vasco. In: EPOFE (ed.) Libro sobre la Pottoka. E-book. <http://www.pottoka.info/lapottoka/libro.php?id=es>.
- Arruga M.V., Intxausti J.I., Monteagudo L.V., Pascual I., Tejedor M.T. (2001) Electrophoretic analysis of four serum protein *loci* in several populations of Pony Vasco-Pottoka. *Arch. Zootec.* **50**, 7–13. (In Spanish).

- Bonnet E., Van de Peer Y. (2002) ZT: a software tool for simple and partial Mantel tests. *J. Stat. Softw.* **7**, 1–12.
- Boyd L., Houpt F.A. (1994) Przewalski's Horse: the History and Biology of an Endangered Species. State University New York Press, Albany, NY.
- 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. *Anim. Genet.* **31**, 39–48.
- Checa M.L. (2004). Análisis de la variabilidad genética en razas equinas autóctonas españolas detectada mediante microsatélites. PhD thesis, Universidad Complutense de Madrid, Madrid Spain. (In Spanish).
- Checa M.L., Vega J.L., García-Atance M.A., Vallejo M., Dunner S. (1998) Genetic variability in two Spanish horse populations: preliminary results. *Arch. Zootec.* **47**, 169–174. (In Spanish).
- Cieslak M., Pruvost M., Benecke N., Hofreiter M., Morales A., Reissmann M., Ludwig A. (2010) Origin and history of mitochondrial DNA lineages in domestic horses. *PLoS ONE* **5**, e15311. doi: 10.1371/journal.pone.0015311.
- Felsenstein J. (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166.
- Goudet J. (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.* **86**, 485–486.
- Harpending H.C., Ward R. (1982) Chemical systematics and human evolution. In: M. Nitecki (ed.), *Biochemical Aspects of Evolutionary Biology*. University of Chicago Press, Chicago, IL, pp. 213–256.
- Hill E.W., Bradley D.G., Al-Barody M., Ertugrul O., Splan R.K., Zakharov I., Cunningham E.P. (2002) History and integrity of thoroughbred dam lines revealed in equine mtDNA variation. *Anim. Genet.* **33**, 287–294.
- Jansen T., Forster P., Levine M.A., Oelke H., Hurles M., Renfrew C., Weber J., Olek K. (2002) Mitochondrial DNA and the origins of the domestic horse. *Proc. Natl Acad. Sci. USA* **99**, 10905–10910.
- Kalinowski S.T., Taper M.L., Marshall T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 1099–1106.
- 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. *Genet. Sel. Evol.* **41**, 5.
- Lindgren G., Backström N., Swinburne J., Hellborg L., Einarsson A., Sandberg K., Cothram G., Vilá C., Binns M., Ellegren H. (2004) Limited number of patrilineages in horse domestication. *Nat. Genet.* **36**, 335–336.
- Lira J., Linderholm A., Olaria C., Durling M.B., Gilbert M.T.P., Ellegren H., Willerslev E., Lidén K., Arsuaga J.L., Götherström A. (2010) Ancient DNA reveals traces of Iberian Neolithic and Bronze Age lineages in modern Iberian horses. *Mol. Ecol.* **19**, 64–78.
- Luis C., Juras R., Oom M.M., Cothram E.G. (2007a) Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. *Anim. Genet.* **38**, 20–27.
- Luis C., Cothram E.G., Oom M.M. (2007b) Inbreeding and genetic structure in the endangered Sorraia horse breed: implications for its conservation and management. *J. Hered.* **98**, 232–237.
- Nei M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.
- Outram A.K., Stear N.A., Bendrey R., Olsen S., Kasparov A., Zaibert V., Thorpe N., Evershed R.P. (2009) The earliest horse harnessing and milking. *Science* **323**, 1332–1335.
- Plante Y., Vega-Pla J.L., Lucas Z., Colling D., De March B., Buchanan F. (2007) Genetic Diversity in a Feral Horse Population from Sable Island, Canada. *J. Hered.* **98**, 594–602.
- Portas M.C.P., Vieira e Brito N., Silva Carvalho I., Vieira Leite J.M. (2001) The conservation of the Garrano horse breed. *Arch. Zootec.* **50**, 171–179. (In Spanish).
- Rendo F., Iriondo M., Jugo B.M., Aguirre A.I., Mazón L.I., Vicario A., Gómez M., Estonba A. (2004) Analysis of the genetic structure of endangered bovine breeds from the Western Pyrenees using DNA microsatellite markers. *Biochem. Genet.* **42**, 99–108.
- Rendo F., Iriondo M., Manzano C., Estonba A. (2009) Identification of horse chestnut coat color genotype using SNaPshot®. *BMC Res. Notes* **2**, 255.
- Reynolds J., Weir B.S., Cockerham C.C. (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**, 767–779.
- Rousset F. (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour* **8**, 103–106.
- Royo L.J., Álvarez I., Beja-Pereira A., Molina A., Fernández I., Jordana J., Gómez E., Gutierrez J.P., Goyache F. (2005) The origins of Iberian horses assessed via mitochondrial ADN. *J. Hered.* **96**, 663–669.
- Royo L.J., Álvarez I., Gutierrez J.P., Fernández I., Goyache F. (2007) Genetic variability in the endangered Asturcón pony assessed using genealogical and molecular information. *Livest. Sci.* **107**, 162–169.
- Solís A. (2005). Markari molekularren bidezko Euskal Autonomi Erkidegoko eta Nafarroako bertako lau zaldi-arrazen karakterizazio genetikoa. PhD Thesis. Universidad del País Vasco/Euskal Herriko Unibertsitatea, Bilbao, Spain. (In Basque).
- Solis A., Jugo B.M., Mériaux J.C., Iriondo M., Mazón L.I., Aguirre A.I., Vicario A., Estonba A. (2005) Genetic diversity within and among four south European native horse breeds based on microsatellite DNA analysis: implications for conservation. *J. Hered.* **96**, 670–678.

- Van Oosterhout C., Hutchinson B., Wills D., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538.
- Vega-Pla J.L., Calderón J., Rodríguez-Gallardo P.P., Martínez 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. *Anim. Genet.* **37**, 571–578.
- Vilà C., Leonard J.A., Götherström A., Marklund S., Sandberg K., Lidén K., Wayne R.K., Ellegren H. (2001) Widespread origins of domestic horse lineages. *Science* **291**, 474–477.
- Wade C.M., Giulotto E., Sigurdsson S., Zoli M., Gnerre S., Imstrand F., Lear T.L., Adelson D.L., Bailey E., Bellone R.R., Blöcker H., Distl O., Edgar R.C., Garber M., Leeb T., Mauceli E., MacLeod J.N., Penedo M.C.T., Raison J.M., Sharpe T., Vogel J., Andersson L., Antczak D.F., Bigagi T., Binns M.M., Chowdhary B.P., Coleman S.J., Della Valle G., Fryc S., Guérin G., Hasegawa T., Hill E.W., Jurka J., Kiiialainen A., Lindgren G., Liu J., Magnani E., Mickelson J.R., Murray J., Nergadze S.G., Onofrio R., Pedroni S., Piras M.F., Raudsepp T., Rocchi M., Røed K.H., Ryder O.A., Searle S., Skow L., Swinburne J.E., Syvänen A.C., Tozaki T., Valberg S.J., Vaudin M., White J.R., Zody M.C., Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, Lander E.S., Lindblad-Toh K. (2009) Genome Sequence, Comparative Analysis, and Population Genetics of the Domestic Horse. *Science* **326**, 865–867.
- Warmuth V., Eriksson A., Bower M.A., Cañón J., Cothran G., Distl O., Glowatzki-Mullis M.L., Hunt H., Luís C., Oom M.M., Tupac-Yupanqui I., Ząbek T., Manica A. (2011) European domestic horses originated in two Holocene refugia. *PLoS ONE* **6**, e18194. doi: 10.1371/journal.pone.0018194.
- Ząbek T., Nogal A., Radko A., Nogal J., Słota E. (2005) Genetic variation of Polish endangered Biłgoraj horses and two common horse breeds in microsatellite *loci*. *J. Appl. Genet.* **46**, 299–305.