sciendo

POPULATION STUDY OF THE PURA RAZA ESPAÑOL HORSE REGARDING ITS COAT COLOUR*

Julia Poyato-Bonilla[1], María José Sánchez-Guerrero[1♦], Rute Dos Santos[2], Mercedes Valera[1]

[1]Departamento de Ciencias Agroforestales, ETSIA, Universidad de Sevilla, Ctra. Utrera km 1, 41013 Sevilla, Spain
[2]ESAE, Instituto Politécnico de Portalegre, 7350-903 Elvas, Portugal | ICAAM – Instituto de Ciências Agrárias e Ambientais Mediterrânicas
♦Corresponding author: v32sagum1@gmail.com

Abstract
Coat colour has always been a valuable trait for horse breeders. However, preferences for this feature have changed over the years. In this research, the Pura Raza Español horse (PRE) population was divided into four subpopulations (Grey, Bay, Black and Others), according to the most frequent coat colours and those of their ancestors. The purpose was to analyse genetic variability, reproductive parameters and distances among subpopulations during three key periods in the history of the breed: before 1960, from 1960 to 2000 and after 2000. The subpopulations composed of animals with ancestors with the same coat colour showed higher values of recent inbreeding (ranging from 7.13% to 10.44%) and a greater Nei's minimum distance between them, as a result of more inbred matings than those carried out in families with members with different coat colours. Non-pure subpopulations also showed more similar recent inbreeding values (between 6.63% and 6.74%). Strikingly, the productive life of Pure bay, Pure black and other subpopulations with minority coat colours was considerably longer (10.79, 10.08 and 9.11 years, respectively) compared to the values of grey PRE horses (6.01 and 7.98 years), which is the subpopulation with the highest census. These results, together with shorter generation intervals of black stallion-offspring (5.51 years via father-son and 6.39 years via father-daughter) and the fact that this coat colour was not present in the breed until two decades ago, highlight the recent trend towards the breeding of black animals.

Key words: equine, genetic diversity, inbreeding, reproductive parameters

The PRE, traditionally known as Andalusian horse, is one of the oldest European horse breeds and the most important horse in Spain according to the census (Sánchez-Guerrero et al., 2016). Its active worldwide population comprises over

200,000 individuals in 2016. In the past, it was widely used as a draft horse for agricultural work and in the army. Nevertheless, due to mechanization of farming in the 1960s and 1970s, the number of PRE used in rural areas declined, while it became increasingly popular as a riding horse. In particular, its use in equestrian sports has been oriented towards the dressage discipline, where it is currently the most important breed participating in Spain (Sánchez et al., 2016). At present, the PRE is also a leisure horse, used under saddle or carriage at popular festivals.

Coat colour is one of the most noteworthy features of an animal and has historically been the object of particular interest to humans (Pruvost et al., 2011). The genes controlling a horse's coat colour have been known for a long time. However, it was only in the last 15 years that functional alleles or marker alleles have been detected at DNA level (Rieder, 2009). A horse's coat colour is mostly controlled by genes at 12 different loci (Thiruvenkadan et al., 2008), and the three basic coat colours of horses are black, bay and chestnut. The genetic control over the basic coat colour of horses resides at two genetic loci, namely Extension (E) and Agouti (A). Among the basic coat colours, bay is dominant to black and both are epistatic to chestnut. The grey coat colour is due to the presence of a dominant gene (G) at the grey locus. Grey is epistatic to all coat colour genes except for white. Thus, a grey horse must have at least one grey parent (Thiruvenkadan et al., 2008).

Horse breeders traditionally associate coat colours with physical qualities and behaviour; for instance, it is a widespread belief that chestnut horses, particularly females, exhibit erratic conduct (Finn et al., 2016). Therefore, a horse's value has always been linked to its coat colouration and pattern; and although some studies support the hypothesis that particular coat colours might be advantageous under certain environmental conditions, genetic studies have shown that the coat colour variation observed in the different horse breeds is the result of domestication and selective breeding (Bellone, 2010; Haase et al., 2007; Horváth et al., 2010; Ludwig et al., 2009).

Currently, according to the PRE studbook, the prevailing coat colour is grey, followed by bay and black. Chestnut, buckskin, roan, isabella, palomino, pearl, overo or dark bay are minority coat colours. Each coat colour owes its origin to the influence of different breeds on PRE: grey originated from animals from northern Africa; bay from the Germanic breeds; chestnut from Arabian horse breeds and black from Holland horses (Andrade, 1954). Nowadays, all coat colours, except those with large white spots on the head and limbs, or of any size on the rest of the body (piebald, pinto or Appaloosa), are accepted in the PRE studbook. The breed standard is dictated by the National Purebred Spanish Horse Breeders' Association (ANCCE, 2017). However, for some years since the foundation of the breed studbook in the 19th century, the registration of horses of some common colours, such as chestnut, has not been permitted. Although grey is the coat colour connected to the origins of the present population and it has always been related to better performance and behaviour, PRE breeders have recently been more inclined to opt for less common colours. These historical differences may involve distinct genetic variability between animals with different coat colours, which is also favoured by the fact that breeders often pair up animals of the same colour group (Bartolomé et al., 2010).

Previous genealogical analyses have shown that there are significant inbreeding levels in this population caused by the low number of breed founders and the selection conducted by breeders for type traits (Valera et al., 2005). In 2003, a PRE Breeding Program was initiated, whose main objectives include improving the breed's morphology, conformation and functionality while maintaining the genetic variability of the breed, by planning matings better (Gómez et al., 2009). It is therefore important to study the population structure periodically, in order to establish correct guidelines for the future management of the breed. The large amount of data registered in the PRE studbook makes it especially reliable for this aim. Moreover, the evolution of coat colour subpopulations throughout the history of the breed has also been researched, dividing the studbook into three periods of time: before 1960, when the PRE was used mainly for agricultural work; from 1960 to 2000: a period of transition where it begins to be used for leisure and equestrian sport as well as remaining linked to the rural environment; and after 2000, from where it is used only for leisure and equestrian sports, mainly dressage.

The objectives of this work were to analyse the genetic diversity of the PRE population, to estimate reproductive parameters that show the breeding strategies used and to establish the genetic distances among the current PRE subpopulations of horses according to their coat colours, based on the genealogical information from their studbooks. This could be a good way to illustrate the mating system used in each subpopulation throughout the breed's history, and to show how animals with certain coat colours have been selected, according to the breeders' preferences.

Material and methods

Pedigree data

Pedigree records from 307,831 horses (148,836 stallions and 158,995 mares) were available from the PRE studbook (Table 1). Each record includes the individual's registration code, as well as their sire and dam, name, birth date and coat colour. In addition, since the 1980s, the pedigree has been recorded through blood-type testing and biochemical polymorphisms, and since 2002 through molecular paternity tests. Previous records, provided by the National Pura Raza Español Breeders' Association (ANCCE), were initially analysed in order to detect and remove possible errors of parenthood assignation.

Four subpopulations of PRE were established, based on phenotypic differences in their coat colour: Grey (n=135,005), Bay (n=96,035), Black (n=28,440) and Others (a subpopulation including those animals with minority coat colours such as chestnut, buckskin, etc., n=7945). Each of the first three subpopulations was divided into two subgroups according to their ancestors' coat colour. One subgroup was composed of animals with at least three generations of ancestors (parents, grandparents and great-grandparents) which had the same coat colour as them:

– Pure grey = animals with at least 3 generations of ancestors with grey coat colour.
– Pure bay = animals with at least 3 generations of ancestors with bay coat colour.
– Pure black = animals with at least 3 generations of ancestors with black coat colour.

– Non-pure grey = at least one ancestor without grey coat colour.
– Non-pure bay= at least one ancestor without bay coat colour.
– Non-pure black = at least one ancestor without black coat colour.

Table 1. Number of animals of each coat colour and percentage by sex included in the PRE studbook per period of time: before 1960, between 1960 and 2000, and after 2000

Period	Grey		Bay		Black		Others	Complete population	Total
	P	NP	P	NP	P	NP			
Number of individuals									
<1960	403	1,750	20*	2,152	0*	310	321	5,956	
1960–2000	25,618	31,500	2,265	15,036	2*	2,037	303	77,662	307,831
>2000	36,932	71,562	10,699	65,863	398	25,693	7,321	224,213	
Males (% of the number of individuals per period)									
<1960	52.61%	44.23%	60.00%*	38.62%	-	22.58%	31.15%	38.13%	
1960–2000	48.86%	48.52%	48.70%	49.76%	100%*	47.96%	43.56%	48.58%	148,836
>2000	49.36%	48.38%	49.11%	49.77%	50%	48.87%	50.61%	48.54%	
Females (% of the number of individuals per period)									
<1960	47.39%	55.77%	40.00%*	61.38%	-	77.42%	68.85%	61.87%	
1960–2000	51.14%	51.48%	51.30%	50.24%	0%*	52.04%	56.44%	51.42%	158,995
>2000	50.64%	51.62%	50.89%	50.23%	50%	51.13%	49.39%	51.46%	

P: pure; NP: non-pure. *Subpopulations with less than 300 animals were not taken into account in this study.

In the Non-pure grey population, 72.26% were the result of matings between grey and bay PRE horses, and the same pattern was found in the Non-pure bay population (39.31%). In contrast, the commonest mating patterns in the Non-pure black subpopulation were between bay and black horses (38.31%), followed by those between two bay horses (20.17%). Animals with an unknown coat colour (0.45%) were not included in any of the previous groups, but were taken into account to estimate the parameters of the whole population.

Furthermore, three periods of time were established in the PRE studbook:
– The first includes animals born from the beginning (1800) of the studbook until 1960.
– The second period encompasses horses born between 1960 and 2000.
– The last includes horses born from 2000 to 2016.

These intervals were established in response to the three main periods of changes in the breed's history: the 1960s was a decade of decline due to its disuse as a draft animal; after this time, the PRE horse was selected primarily for its beauty and racial pattern. It was not until the beginning of the 21st century that its use as a sport horse was promoted. Important fluctuations in the number of animals over time have occurred in all coat colour subpopulations. The data of subpopulations with fewer than 300 individuals are displayed for informative purposes (Pure bay before 1960, Pure black before 1960 and Pure black between 1960 and 2000), but these subpopulations were not included in this research because they were not considered sufficiently relevant.

Methodology

Reproductive parameters

In order to characterize the possible differences in PRE breeding strategies depending on coat colour over time, certain reproductive parameters were estimated. 'Age at first and last offspring' is the age (in years) at which sires and dams had their first and last son or daughter, which shows the length of time individuals are kept as breeding animals and the age up to which they are used for that purpose. In addition, the period of time during which the individuals were kept as breeding animals (productive life) was calculated as the difference in years between the birth date of an animal's last offspring and the birth date of its first one. The total number of sons and daughters of stallions and mares, or number of offspring per breeding animal, gives us insight into the permissiveness in spreading the genes of certain individuals. These parameters were only computed for currently non-active individuals, using two criteria: 1) horses whose reproductive age has ended or horses over 30 years old (an age at which it is estimated that the horse is physiologically unable to procreate), 2) stallions and mares which have had no offspring in the previous four years.

Generation intervals were computed as the average age of parents at the birth of their progeny which were kept for reproduction, for the 4 selection paths (father–son, father–daughter, mother–son and mother–daughter). As inbreeding and pedigree depth are inter-dependent, pedigree completeness was assessed by the equivalent complete generations. Equivalent complete generation (EqG) was computed as the addition, for all known ancestors, of the terms computed as the sum of $(\frac{1}{2})^n$, where n is the number of generations separating the individual from each known ancestor (Maignel et al., 1996). Genealogical analyses were computed using the program Endog v4.8 (Gutiérrez and Goyache, 2005).

Genetic variability and probability of gene origin

The genetic variability was studied across the effective number of founders (f_e), effective number of ancestors (f_a), founder genome equivalents (f_g) and the number of ancestors that account for 50% of the gene diversity of the whole population ($N_{50\%}$). f_e (Lacy, 1989) is the reciprocal of the probability that two alleles selected at random in the studied population originated from the same founder. This parameter reflects the number of animals which, given equal contributions, could produce the same degree of genetic variability found in the population studied. f_a is defined as the minimum number of ancestors, not necessarily founders, which account for the complete genetic diversity of a population (Boichard et al., 1997). This term considers the diversity loss caused by the unbalanced use of individuals, which, in turn, leads to bottlenecks, which are the major cause of gene loss in certain populations. f_g (Ballou and Lacy, 1995) is defined as the number of founders that would be expected to produce the same genetic diversity as in the population in study if the founders were equally represented and no loss of alleles had occurred. It was obtained by the inverse of twice the average coancestry of the individuals included in a pre-defined reference population, following Caballero and Toro, 2000. The percentage of the

common founders of subpopulations was also estimated in order to assess the original closeness of the coat colour subpopulations.

In addition, Genetic Conservation Index (*GCI*; Alderson, 1992) was used to assess the genetic contribution of founders for each animal which correspond with the initial genetic information of the breed. This index is estimated as: $GCI = \dfrac{1}{\Sigma p^2{}_i}$, with p_i the proportion of genes of founder *i* in the genome of an individual.

The proportional contribution of each coat colour subpopulation to the total diversity was characterized by the loss or gain of genetic diversity in the whole population after removal of the subpopulation from the complete population (Caballero and Toro, 2002). The Remaining Gene Diversity (*GD*) is the Nei's gene diversity remaining after the elimination of the corresponding subpopulation.

The probability that an individual possesses two identical alleles by descent at a randomly chosen locus can be measured by the inbreeding coefficient (*F*) (Málecot, 1948) and was computed following Meuwissen and Luo (1992). At population level, *F* represents the percentage of heterozygous loci in the original population that have become homozygous in the present population due to related matings.

To study the representation of an animal in the whole pedigree, independently of the knowledge of its own pedigree, the average relatedness coefficient (*AR*) was computed. *AR* is the probability that an allele randomly chosen from the population in study belongs to an individual. It is computed as the average of the coefficients integrating the row from the individual in the numerator relationship matrix and it takes into account, simultaneously, the inbreeding and coancestry coefficients, so it can be used as a measure of the inbreeding of the whole or sub-population (Goyache et al., 2003; Gutiérrez et al., 2003). Both *F* and *AR* can be used alternatively or complementarily to predict the long-term inbreeding of a population, as *AR* takes into account the percentage of the complete pedigree originating from a given founder at population level (Gutiérrez and Goyache, 2005).

Distances between subpopulations

Genetic distances between subpopulations were measured by Nei's minimum distance (Nei, 1987), following Caballero and Toro (2000, 2002), as $D_{ij} = [(f_{ii}+f_{jj})/2] - f_{ij}$, with f_{ij} the average pairwise coancestry coefficient among individuals of two subpopulations i and j; and f_{ii} and f_{jj} the within-subpopulation coancestry using the program Endog v4.8 (Gutiérrez and Goyache, 2005).

Results

The PRE population census for the three periods is shown in Table 1. On the whole, grey subpopulations have always had the largest number of animals, while other subpopulations have had less. Individually, the lowest number of animals corresponds to the set of Pure black, which also coincides with the subpopulation with the greatest increase in individuals (2 in 1960–2000 and 398 in 2000–2016). The percentages of males and females are balanced, with a slightly higher average percentage of females than males (51.65% and 48.35%).

Table 2. Reproductive parameters. Average age at first and last offspring (separated by a dash) in years, average productive life of breeding animals in years and average number of offspring per breeding animal in each subpopulation by coat colour and period of time

		Age at first-last offspring (years)			Productive life (years)			Number of offspring/breeding animal		
		<1960	1960–2000	>2000	<1960	1960–2000	>2000	<1960	1960–2000	>2000
Grey	P	6.93–14.11	5.87–12.84	6.07–13.92	7.40	7.10	7.98	3.08	2.52	4.46
	NP	7.15–13.63	6.01–13.70	5.72–11.64	6.75	7.81	6.01	2.47	2.90	4.19
Bay	P	–	5.66–13.43	6.00–15.22	–	10.46	10.79	–	3.89	2.97
	NP	7.79–13.01	5.77–13.38	5.67–12.49	5.54	10.11	6.93	1.45	4.09	3.69
Black	P	–	–	6.91–16.89	–	–	10.08	–	–	2.06
	NP	8.24–12.51	5.31–13.20	5.68–13.82	4.66	11.64	8.11	0.49	6.91	2.89
Others	NP	8.07–12.37	7.72–13.84	4.72–14.50	4.72	8.18	9.11	0.47	3.45	2.50
Complete population		11.23–16.85	5.90–13.33	5.11–10.61	6.23	7.56	7.24	2.63	5.12	3.86

P: pure; NP: non-pure.

Table 3. Generation intervals computed for the four pathways (father–son, father–daughter, mother–son and mother–daughter) for the coat colour subpopulations and the complete population, before 1960, between 1960 and 2000 and after 2016

		Generation intervals											
		Father–son			Father–daughter			Mother–son			Mother–daughter		
		<1960	1960–2000	>2000	<1960	1960–2000	>2000	<1960	1960–2000	>2000	<1960	1960–2000	>2000
Grey	P	9.25	9.99	10.50	10.21	10.03	10.86	9.71	9.35	9.53	9.74	9.15	9.28
	NP	11.83	10.82	9.94	12.60	10.71	9.95	13.28	10.00	9.26	12.30	9.76	8.96
Bay	P	–	9.91	10.26	–	10.14	10.14	–	9.69	8.77	–	9.42	8.60
	NP	12.35	10.59	9.41	13.29	10.37	9.46	13.72	9.64	8.75	12.23	9.38	8.58
Black	P	–	–	5.51	–	–	6.39	–	–	9.16	–	–	8.43
	NP	9.10	9.45	8.81	13.04	9.82	8.73	9.62	9.16	8.21	12.64	8.87	8.15
Others	NP	9.37	10.29	9.70	10.04	9.86	8.87	10.57	9.22	8.57	11.23	9.11	8.76
Complete population		16.64	10.40	9.54	18.79	10.38	9.78	15.38	9.64	8.82	18.42	9.40	8.76

P: pure; NP: non-pure.

Table 4. Pedigree completeness, probability of gene origin and genealogical parameters. Equivalent complete generations (EqG), number of founders (f_e), number of ancestors (f_a), founder genome equivalents (f_g), number of ancestors that explain the 50% of the gene diversity (N_{50}), Genetic Conservation Index (GCI), and remaining gene diversity (GD). Inbreeding coefficient (F) and Average Relatedness (AR) are shown for each period of time: before 1960, from 1960 until 2000 and after 2000.

	Grey		Bay		Black		Others	Complete population
	P	NP	P	NP	P	NP		
EqG	9.17	8.96	9.33	9.44	10.65	8.90	9.88	9.08
f_e	15	26	49	37	35	37	31	28
f_a	13	20	20	21	22	21	22	19
f_g	8.35	8.99	7.18	8.75	6.68	8.77	10.09	115.22
$N_{50\%}$	5	7	8	8	9	8	8	7
GCI	9.98	12.32	12.07	11.71	9.63	11.71	13.05	11.14
GD	0.95	0.95	0.94	0.94	0.94	0.94	0.94	-
Inbreeding coefficient (*F*)								
Period								
<1960	11.09%	5.31%	-	2.45%	-	1.74%	1.44%	3.66%
1960–2000	11.46%	7.93%	7.79%	6.94%	-	6.92%	7.80%	8.53%
>2000	10.44%	6.73%	7.13%	6.63%	9.93%	6.74%	7.33%	7.13%
Average Relatendness (*AR*)								
<1960	14.19%	6.84%	-	3.04%	-	1.51%	1.89%	11.00%
1960–2000	17.21%	12.19%	9.76%	10.55%	-	10.59%	11.18%	11.98%
>2000	16.14%	11.30%	9.92%	10.68%	9.40%	10.67%	11.17%	11.00%

P: pure; NP: non-pure.

Table 5. Above the diagonal: matrix of percentage of common founders between subpopulations. Below the diagonal: matrix of Nei's minimum distances

D_{ij} / % common founders	P grey	NP grey	P bay	NP bay	P black	NP black	Other coat colours
P grey	-	12.31%	12.53%	9.63%	11.52%	11.90%	11.91%
NP grey	0.22	-	5.42%	7.41%	4.77%	6.45%	6.32%
P bay	2.09	1.10	-	22.99%	46.16%	35.86%	31.16%
NP bay	0.52	0.08	0.64	-	20.59%	27.41%	23.00%
P black	1.18	0.73	1.43	0.65	-	33.68%	29.70%
NP black	0.62	0.15	0.64	0.05	0.49	-	29.28%
Other coat colours	0.61	0.13	0.67	0.05	0.72	0.11	-

P: pure; NP: non-pure.

The reproductive parameters are shown in Table 2: age at first offspring ranged from 4.72 years in the most recent 'Others' subpopulation to 11.23 years in the complete population before 1960. Interestingly, the age at last offspring varied between an average of 10.61 years in the complete population and 16.89 in the pure black subpopulation from 2000 to now. The shortest productive lives in the PRE were those of Non-pure black animals (4.66 years) and the longest, 11.64 years, for the same subpopulation during the first and second periods of time, respectively. The number of offspring per breeding animal underwent substantial changes through the periods of time: with the exception of the grey and Pure black lines (where the opposite situation occurs), the number of progeny in the rest of the subpopulations fell from the second to the last period. Notably, Non-pure black animals had an average of 6.91 offspring between the years 1960 and 2000. Generation intervals, computed for the four pathways and their evolution throughout time are displayed in Table 3. The father-offspring pathways were the longest in all the subpopulations and periods of time, while the mother-offspring were the shortest. All the subpopulations, except the Pure grey, have experienced a decline from the past until now.

The Pedigree completeness of the PRE population, shown in Table 4, resulted in 9.08 equivalent generations, with a minimum of 8.90 and a maximum of 10.65 from Non-pure and Pure black subpopulations, respectively. The subpopulations with the least diverse probability of gene origin (f_e, f_a and N_{50}) were the grey ones, and the lowest number of founder genomes corresponded to the Pure black subpopulation. The Genetic Conservation Index (*GCI*) ranged from 9.63 in the case of Pure black to 13.05, from the 'Others' group. All the subpopulations shared a similar average for Remaining Gene Diversity (*GD*), which was 0.945. The maximum inbreeding (F) and average Relatedness (AR) coefficients corresponded to the Pure grey subpopulation and minimum to all the Non-pure subpopulations. In all cases, there was a rise in F and AR between the first period of time and the second, which later fell.

Regarding between-subpopulation Nei's minimum distances, the most distant subpopulations were the pure coat colour subpopulations. In contrast, minimum distances were found between the Non-pure bay and Non-pure black, Non-pure bay and Others, and Non-pure bay and Non-pure grey subpopulations. All distances are shown in Table 5, along with the percentage of shared founders. The grey subpopulations were the ones that shared fewer founders with other coat colours (only 4.77% to 7.41% of founders were common to Non-pure grey and the rest of the coat colour subpopulations except Pure grey, while Pure bay and Pure black shared 46.6% of founders.

Discussion

While many species of wild animals are relatively uniform in colour, domestic stock, such as horses, show a wide variety of coat colour patterns (Rieder, 2009). Human preferences and demand may have favoured rare alleles or led to previously

unknown phenotypes due to selective breeding (Klungland and Vage, 2000). The selection of coat colour as a division criterion responds to the fact that for many years, market (and hence breeders') preferences have tended more towards a certain colour or another, either due to the attribution of better performance or for their scarcity, as seen in breeds such as Hanoverian (Icken et al., 2007).

An innate preference for white variants appears to be a universal human trait, and white horses have been prized for at least the last 2500 years (Linderholm and Larson, 2013). The prestige of riding a white horse has thus led to selection by humans of the grey-causing mutation; this mutation is by far the most common cause of white colour in horses (Rosengren-Pielberg et al., 2008). Since Grey is a dominant gene (Rieder, 2009) and it has been the majority coat colour since the foundation of the PRE studbook, the breed could follow the steps of Lipizzan and become an almost totally grey population unless selection to obtain other coat colours is carried out. Until the 18th century, Lipizzans had a wide variety of coat colours (including dun, bay, chestnut, black, piebald and skewbald). Grey was deliberately selected as a desirable feature, and so became the coat colour of the overwhelming majority of Lipizzan horses (Swinney, 2010). Fortunately, from the census point of view, the PRE population has undergone substantial changes (Table 1). Even if Grey subpopulations comprise the largest number of individuals, they have experienced the smallest increases (of 1.44 and 2.27-fold) from the second to the third period. This clearly indicates a selection towards other coat colours as the way to avoid the predominance of grey and increase the number of individuals with different coat colours. Pure black animals were notably not present in the breed until the second time period, when there were only 2 individuals. Notwithstanding, from 2000 to 2016, its subpopulation comprised 398 horses, which represents an increase of almost 200-fold. Meanwhile, Non-pure black animals have always been present in the PRE, but have also undergone a remarkable expansion. This suggests that many black animals were not the result of matings aimed at specifically obtaining that colour but were, in fact, the offspring of horses of other coat colours, since black animals can be obtained from the mating of animals of very different coat colours. Theoretically, when two bay horses are mated, there is a 6% probability of them producing a black foal, while the probability between bay and black horses is 56.26%. In our population, Non-pure black horses mainly descend from two bay horses (probably due to the frequency of this combination) and bay and black horses (probably due to the high probability of this producing a black foal). In fact, in ancient times, black was an undesirable coat colour, since black was associated with the horses that pulled hearses. However, this belief was replaced by the idea that animals of this coat colour were stronger, more spirited and also hid morphological defects better. Furthermore, the increase of the black subpopulations could be indirectly related to the acceptance in 2004 of the chestnut coat colour as a breed standard. Thus, matings between two black individuals would have been undesirable in order to avoid a possible chestnut offspring for years. Besides, even though the Pure bay subpopulation in the early period of the breed consisted of only 20 horses, this subpopulation reached 2,265 individuals in 1960–2000 (an increase of 113.25-fold), implying that this coat colour did not become popular until the second time period.

From the analysis of the reproductive parameters (Table 2), we can observe that coat colour shows the different reproductive strategies of breeders. Regarding the age at first offspring, Non-pure subpopulations and minority coat colours (Others) have their progeny sooner than the rest. On the other hand, the age at last offspring is considerably higher, not only in the case of Pure bay and Pure black, but also in Others, which demonstrates that these animals are kept for reproduction for longer periods, possibly with the aim of producing a greater number of offspring.

'Productive life' is defined as the years that animals are active as breeding animals. As the date of neither death nor retirement is available for all animals in the PRE studbook, this parameter has been estimated from the difference between the age at which animals have their last and first offspring. Interestingly, all PRE, regardless of their coat colour, experienced an increase in their productive life over the period between 1960 and 2000. This rise in productive life coincides with the decline of the use of the breed as a draft horse and its rise in popularity as a pleasure or sport horse. The beginning of the 21st century has seen an increase in the productive life of Pure subpopulations and Others, but a fall in Non-pure ones, which may be due to a change in breeders' policies towards the production of horses with a greater probability of producing offspring with certain coat colours.

The number of offspring per breeding animal provides insight into the importance of coat colour in relation to the number of descendants they leave. The results show that PRE had the fewest offspring before 1960, while the Grey subpopulations made a greater contribution of descendants per stallion or mare. In the final period, the Grey subpopulations were the most prolific, with more offspring than the average of the complete population, followed by Non-pure bay. The least prolific subpopulation since 2000 was Pure black. The number of offspring/breeding animal for the complete population (3.86) was similar to the mean value of offspring registered in Lusitano mares: 4.00 (Vicente et al., 2012).

The generation interval (Table 3) is the amount of time required to replace one generation with the next. This parameter is inversely proportional to the selection response and is important in the study of genetic structures of populations, due to its relation with losses in genetic variability. In this study, generation intervals of PRE were lower for the mother-offspring pathways, except for the Pure black subpopulation. Interestingly, Pure black stallions have their first offspring very young (5.51 years for sons and 6.39 years for daughters). The small number of Pure black stallions means that these individuals are chosen as breeding animals at very early ages. The other coat colour subpopulations showed more homogeneous generational intervals, with the grey subpopulations the longest. The evolution of the breed throughout the three periods of time showed a remarkable decline in generation intervals. The average PRE generation intervals were 9.78 years, in line with those reported in Lusitano (10.3 years; Vicente et al., 2012), Hanoverian (10; Hamann and Distl, 2008), Spanish Arab-derived horses (10.1–12.0; Cervantes et al., 2009), Mangalarga (9.49; Mota et al., 2006) and Trakehner (10.2; Teegen et al., 2009); nonetheless, the PRE generation interval was higher than that reported in a draught breed like Franches-Montagnes (8.9) (Hasler et al., 2011).

The Pedigree completeness, probability of gene origin and genealogical parameters are shown in Table 4. The PRE is a closed population with a well-established studbook going back many years, which makes calculations of parameters more reliable. The number of equivalent generations in PRE is approximately 9, which is a high value compared to those of other pedigrees: 3.19 (Menorca Horse; Solé et al., 2014), 4.43 (German Paint Horse; Siderits et al., 2013) or 7.9 (Spanish Arab Horse; Cervantes et al., 2008 a). However, other breeds such as the Lipizzan (Zechner et al., 2002) have higher values of equivalent generations (15.2). Pedigree completeness did not differ much among coat colour subpopulations in the PRE, which proves the uniformity of the data collection, regardless of the colour of the animal.

The effective number of founders, defined as the number of equally-contributing founders resulting in the same genetic diversity as found in the reference population, was 28, although it differed somewhat among coat colour subpopulations (standardised f_e values by subpopulation size show that Pure black had the highest number in contrast to the Pure grey subpopulation, which had the lowest). In contrast, the numbers reported for the Spanish Arab horse, 39.5 (Cervantes et al., 2008) and Lipizzan, 48.2 (Zechner et al., 2002) were considerably higher. Founders contributing to more than one subpopulation can explain the genetic origin and distances between them. Our results have shown that subpopulations sharing the most founders were Pure bay and Pure and Non-pure black, followed by 'Others'. Conversely, the largest subpopulation, Non-pure grey, had no more than 7.41% of founders in common with the rest of the subpopulations. This suggests that, at first, subpopulations with less common coat colours were established from Non-grey animals, possibly to avoid this colour, due to its dominance, while founders of Grey subpopulations were used more exclusively to obtain grey offspring. The effective number of ancestors can be used to identify bottlenecks. The number for the f_a PRE population, 19, was within the values of Lipizzan and Spanish Arab (Zechner et al., 2002; Cervantes et al., 2008 b), which ranged between 16.5 and 29.3. However, it was lower than the 77.7 registered for the German Hanoverian Horse (Hamann and Distl, 2008) and the figure of around 200 recorded in the German Paint Horse (Siderits et al., 2013).

The number of ancestors that account for 50 per cent of the genetic variability of the population varied from 5 to 9, in keeping with the subpopulations which showed higher and lower values of f_e and f_a. The GCI for the complete population indicated that an average animal has a mean contribution of 11.14 founders, while, depending on the coat colour, this value can be as high as 13.05 (Others). In the case of Lusitano horse, the GCI for a reference population was lower than the PRE population: 9.5 (Vicente et al., 2012).

As regards the remaining gene diversity after removing a subpopulation, all the coat colour subpopulations contributed with almost the same gene diversity to the complete population. Nevertheless, despite being of the same order, the Pure grey subpopulation provided less genetic diversity to the rest of the population. Low *GD* values show that the loss of any of the subpopulations does not imply a great loss of genetic variability. One interpretation of this fact is that all subpopulations are essential for the total genetic diversity at the same level, and it is therefore important to maintain the current coat colour proportions in the PRE and to avoid a disproportion-

ate increase in the homozygous Pure grey subpopulation. This work could be used to take decisions for a future PRE germplasm bank and genomic selection, since there might be a risk that, due to the high number of grey PRE horses (46.86% of the whole population), the majority of the horses selected by breeders were of this colour. For this reason, it would be essential to highlight the importance of maintaining, at least, the current proportions of coat colours.

The genetic relevance of the diversity of a subpopulation within a population is an important subject for study, using genealogical and molecular data which is then easily computed following the methodology of Caballero and Toro (2002), based on the coancestry matrix and the comparison of between- and within-coancestry coefficients.

In this study, the mean inbreeding of all coat colour subpopulations increased from the first time period of the studbook until 1960–2000, with a slight decrease in the final years. Pure colour subpopulations show higher inbreeding coefficients than Non-pure ones, since these subpopulations are based on matings of more closely related animals. The study of this breed by Valera et al. (2005), which only included information up to 1998, found an inbreeding value of 8.48%. This data is in line with the 8.53% obtained in our results in animals born from 1960 to 2000. The inbreeding coefficient of the complete PRE population from 2000 up to the present was 7.13%, which demonstrates that the genetic improvement programme and mating policies are working well. In comparison, the inbreeding coefficient for Lusitano horses was 11.34% (Vicente et al., 2012), for Lipizzan, 10.8% (Zechner et al., 2002) and for Spanish Arab, 7.0% (9.8% for individuals born 1995–2004; Cervantes et al., 2008 b).

The average relatedness of the complete PRE population oscillated around 11.00% and 11.98% during the three established periods of time. According to the study by Valera et al. (2005), this parameter was 12.25% in animals born from the foundation of the studbook until 1998. In comparison, AR in the Spanish Arab Horse (Cervantes et al., 2008 b) was 9.1% and 11.64% in the Lusitano breed (Vicente et al., 2012). However, notable differences between coat colours subpopulations can be observed. The higher AR values belong to Pure grey animals (16.14% in the final period of time), followed by Non-pure grey (11.30%). The Pure black subpopulation had the lowest AR (9.40%).

Nei's minimum genetic distance (Table 5) is a tool to investigate population subdivision. The results showed that there is a lack of differentiation between coat colour subpopulations in this breed. However, greater distances can be observed among Pure colour subpopulations, as a consequence of more related matings among individuals and, therefore, higher coancestry values. These results reinforce the previous study by Bartolomé et al. (2010), which accounts for the minor genetic distances between coat colour subpopulations due to their close genetic relation. Also, in the case of the PRE breed, colour phenotypes may have played a major role during early domestication events and initial selection (Andersson and Georges, 2004). As reflected by the percentage of founders in common (Table 5), a large number of shared founders does not imply greater distances among subpopulations. Interestingly, one of the most distant pair of subpopulations, Pure bay and Pure black, share the largest number of founders (46.16%). This demonstrates that, even if subpopulations had

a common genetic origin, they have followed a differentiation process based on selection, mainly due to the search for animals with the desired coat colours.

The results obtained in this study highlight the fact that PRE breeders' appreciation for coat colours has varied over time, thus modifying mating systems and, as a consequence, inbreeding and relatedness between animals. In addition, changes in the breed studbook to allow or prohibit the registration of horses of certain coat colours and the fact that the commonest current coat colour (grey) is epistatic over the rest, accentuate the importance of maintaining all coat colours. All of these findings strongly suggest that it is crucial to encourage matings between PRE horses of different coat colours, in order to act as gene reservoirs and to preserve the genetic variability. However, the direct practical implementation of this policy is not always straightforward, as the preferential use of certain stallions and the covering fee also play key roles in mating decisions. Moreover, the breeding programme for the PRE prioritises the selection of animals based on their performance while maintaining the levels of genetic variability. We hope, therefore, that these results could help to influence the decision to maintain all the present coat colours in current breeding programmes.

Acknowledgements

References

A l d e r s o n G.L.H. (1992). A system to maximize the maintenance of genetic variability in small populations. Genet. Conserv. Domest. Livest., 2: 18–29.

ANCCE (2012). Specific Purebred Spanish Horse Stud Book Rules and Regulations (BOE nº 127, May 28, 2012 [WWW Document]. URL http://www.ancce.es/ver_contenido.php?id=TXT_DATOS&menu=1&i=en

A n d e r s s o n L., G e o r g e s M. (2004). Domestic-animal genomics: Deciphering the genetics of complex traits. Nat. Rev. Genet., 5; doi:10.1038/nrg1294.

A n d r a d e R. de (1954). Deberá ser castaño el caballo español? In: Alrededor Del Caballo Español: Colección de Estudios. Lisboa: Sociedade Astória, pp. 431–445.

B a l l o u J.D., L a c y R.C. (1995). Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Population management for survival and recovery: analytical methods and strategies in small population management, Ballou J.D., Gilpin M., Foose T.J. (eds). New York: Columbia University Press, pp. 76–111.

B a r t o l o m é E., G o y a c h e F., M o l i n a A., C e r v a n t e s I., V a l e r a M., G u t i é r r e z J.P. (2010). Pedigree estimation of the (sub) population contribution to the total gene diversity: the horse coat colour case. Animal, 4: 867.

B e l l o n e R.R. (2010). Pleiotropic effects of pigmentation genes in horses. Anim. Genet., 41: 100–110.

B o i c h a r d D., M a i g n e l L., V e r r i e r E. (1997). The value of using probabilities of gene origin to measure genetic variability in a population. Genet. Sel. Evol., 29: 5–23.

C a b a l l e r o A., T o r o M.A. (2002). Analysis of genetic diversity for the management of conserved subdivided populations. Conserv. Genet., 3: 289–299.

C a b a l l e r o A., T o r o M.A. (2000). Interrelations between effective population size and other pedigree tools for the management of conserved populations. Genet. Res., 75: 331–343.

C e r v a n t e s F., Gutiérrez J.P., Fernández I., Gomez E., Alvarez I., Díez J., Royo L.J. (2003). Using pedigree information to monitor genetic variability of endangered populations: The Xalda sheep breed of Asturias as an example. J. Anim. Breed. Genet., 120: 95–105.

C e r v a n t e s I., G o y a c h e F., Molina A., V a l e r a M., Gutiérrez J.P. (2008 a). Application of individual increase in inbreeding to estimate realized effective sizes from real pedigrees. J. Anim. Breed. Genet., 125: 301–310.

C e r v a n t e s I., M o l i n a A., C e r v a n t e s F., G u t i é r r e z J.P., V a l e r a M. (2008 b). Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. Livest. Sci., 113: 24–33.

C e r v a n t e s I., G u t i é r r e z J.P., M o l i n a A., C e r v a n t e s F., V a l e r a M. (2009). Genealogical analyses in open populations: The case of three Arab-derived Spanish horse breeds. J. Anim. Breed. Genet., 126: 335–347.

F i n n J.L., H a a s e B., W i l l e t C.E., v a n R o o y D., C h e w T., Wade C.M., H a m i l t o n N.A., V e l i e B.D. (2016). The relationship between coat colour phenotype and equine behaviour: A pilot study. Appl. Anim. Behav. Sci., 174: 66–69.

G ó m e z M.D., V a l e r a M., M o l i n a A., G u t i é r r e z J.P., C e r v a n t e s F. (2009). Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. Livest. Sci., 122: 149–155.

G o y a c h e F., G u t i é r r e z J.P., F e r n á n d e z I., G o m e z E., A l v a r e z I., D í e z J., R o y o L.J., (2003). Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. J. Anim. Breed. Genet., 120: 95–105.

G u t i é r r e z J.P., C e r v a n t e s F. (2005). A note on ENDOG: A computer program for analysing pedigree information. J. Anim. Breed. Genet., 122: 172–176.

G u t i é r r e z J.P., G o y a c h e F. (2005). A note on ENDOG: A computer program for analysing pedigree information. J. Anim. Breed. Genet., 122: 172–176.

G u t i é r r e z J.P., A l t a r r i b a J., D í a z C., Q u i n t a n i l l a R., C a ñ ó n J., P i e d r a f i t a J. (2003). Pedigree analysis of eight Spanish beef cattle breeds. Genet. Sel. Evol., 35: 43–64.

H a a s e B., B r o o k s S.A., S c h l u m b a u m A., A z o r P.J., B a i l e y E., A l a e d d i n e F., M e v - i s s e n M., B u r g e r D., P o n c e t P.A., R i e d e r S., L e e b T. (2007). Allelic heterogeneity at the equine KIT locus in dominant white (W) horses. PLoS Genet., 3: 2101–2108.

H a m a n n H., D i s t l O. (2008). Genetic variability in Hanoverian warmblood horses using pedigree analysis. J. Anim. Sci., 86: 1503–1513.

H a s l e r H., F l u r y C., M e n e t S., H a a s e B., L e e b T., S i m i a n e r H., P o n c e t P.A., R i e - d e r S. (2011). Genetic diversity in an indigenous horse breed – implications for mating strategies and the control of future inbreeding. J. Anim. Breed. Genet., 128: 394–406.

H o r v á t h G., B l a h ó M., K r i s k a G., H e g e d ü s R., G e r i c s B., F a r k a s R., A k e s s o n S. (2010). An unexpected advantage of whiteness in horses: the most horsefly-proof horse has a depolarizing white coat. Proc. Biol. Sci., 277: 1643–1650.

I c k e n W., B e n n e w i t z J., K a l m E. (2007). Analysis of auction data for horses and influence factors on pricing. Zuchtungskunde, 79: 111–118.

K l u n g l a n d H., V a g e D.I. (2000). Molecular genetics of pigmentation in domestic animals. Curr. Genomics, 1.

L a c y R.C. (1989). Analysis of founder representation in pedigrees: founder equivalent and founder genome equivalents. Zoo. Biol. 8: 111–123.

L i n d e r h o l m A., L a r s o n G. (2013). The role of humans in facilitating and sustaining coat colour variation in domestic animals. Seminars in Cell and Developmental Biology. Elsevier Ltd., 24: 587–593.

L u d w i g A., P r u v o s t M., R e i s s m a n n M., B e n e c k e N., B r o c k m a n n G.A., C a s t a n o s P., C i e s l a k M., L i p p o l d S., L l o r e n t e L., M a l a s p i n a s A.-S., S l a t k i n M., H o f r e - i t e r M. (2009). Coat color variation at the beginning of horse domestication. Science, 324: 485–485.

M a i g n e l L., B o i c h a r d D., V e r r i e r E. (1996). Genetic variability of French dairy breeds estimated from pedigree information. Interbull, 14: 49–54.

M á l e c o t L. (1948). Les Mathématiques de l'Hérédité. Masson et Cie, Paris, France.

M e u w i s s e n T.H.E., L u o Z. (1992). Computing inbreeding coefficients in large populations. Genet. Sel. Evol., 24: 305–313.

M o t a M.D.S., P r a d o R.S., S o b r e i r o J. (2006). Caracterização da população de cavalos Mang-alarga no Brasil. Archivos de Zootecnia, 55: 31–37.

N e i M. (1987). Molecular Evolutionary Genetics. Columbia University Press, New York.

P r u v o s t M., B e l l o n e R., B e n e c k e N., S a n d o v a l - C a s t e l l a n o s E., C i e s l a k M. (2011). Genotypes of predomestic horses match phenotypes painted in Paleolithic works of cave art. Proc. Natl. Acad. Sci., 108: 18626–18360.

R i e d e r S. (2009). Molecular tests for coat colours in horses. J. Anim. Breed. Genet., 126; doi:10.1111/j.1439-0388.2009.00832.x

S á n c h e z - G u e r r e r o M.J., M o l i n a A., G ó m e z M.D., P e ñ a F., V a l e r a M. (2016). Relationship between morphology and performance: Signature of mass-selection in Pura Raza Español horse. Livest. Sci., 185: 148–155.

S á n c h e z M.J., B a r t o l o m é E., V a l e r a M. (2016). Genetic study of stress assessed with infrared thermography during dressage competitions in the Pura Raza Español horse. Appl. Anim. Behav. Sci., 174: 58–65.

S i d e r i t s M., B a u m u n g R., F u e r s t - W a l t l B. (2013). Pedigree analysis in the German Paint Horse: Genetic variability and the influence of pedigree quality. Livest. Sci., 151: 152–157.

S o l e M., C e r v a n t e s I., G u t i e r r e z J.P., G o m e z M.D., V a l e r a M. (2014). Estimation of genetic parameters for morphological and functional traits in a Menorca horse population. Spanish J. Agric. Res., 12: 125–132.

S w i n n e y N.J. (2006). Horse Breeds of the World. Hamlyn, London.

T e e g e n R., E d e l C., T h a l l e r G. (2009). Population structure of the Trakehner Horse breed. Animal, 3: 6–15.

T h i r u v e n k a d a n A.K., K a n d a s a m y N., P a n n e e r s e l v a m S. (2008). Coat colour inheritance in horses. Livest. Sci., 117: 109–129.

V a l e r a M., M o l i n a A., G u t i é r r e z J.P., G ó m e z J., C e r v a n t e s F. (2005). Pedigree analysis in the Andalusian horse: Population structure, genetic variability and influence of the Carthusian strain. Livest. Prod. Sci., 95: 57–66.

V i c e n t e A.A., C a r o l i n o N., G a m a L.T. (2012). Genetic diversity in the Lusitano horse breed assessed by pedigree analysis. Livest. Sci., 148: 16–25.

Z e c h n e r P., S ö l k n e r J., B o d o I., D r u m l T., B a u m u n g , R., A c h m a n n R., M a r t i E., H a b e F., B r e m G. (2002). Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. Livest. Prod. Sci., 77: 137–146.