Animal (2010), **4:6**, pp 867–875 © The Animal Consortium 2010
doi:10.1017/S1751731110000182

animal

Pedigree estimation of the (sub) population contribution to the total gene diversity: the horse coat colour case

E. Bartolomé[1†], F. Goyache[2], A. Molina[3], I. Cervantes[4], M. Valera[1] and J. P. Gutiérrez[4]

[1]Departamento de Ciencias Agroforestales, EUITA, Universidad de Sevilla, Ctra. Utrera, km1, 41013, Sevilla, Spain; [2]Área de Genética y Reproducción Animal, SERIDA-Deva, Camino de Rioseco, 1225, E-33394, Gijón (Asturias), Spain; [3]Departamento de Genética, Facultad de Veterinaria, Universidad de Córdoba, Ctra. Madrid-Córdoba, km396a, 14071, Córdoba, Spain; [4]Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense de Madrid, Avda. Puerta de Hierro s/n, 28040, Madrid, Spain

(Received 1 September 2009; Accepted 15 December 2009; First published online 22 February 2010)

A method to quantify the contribution of subpopulations to genetic diversity in the whole population was assessed using pedigree information. The standardization of between- and within-subpopulation mean coancestries was developed to account for the different coat colour subpopulation sizes in the Spanish Purebred (SPB) horse population. The data included 166 264 horses registered in the SPB Studbook. Animals born in the past 11 years (1996 to 2006) were selected as the 'reference population' and were grouped according to coat colour into eight subpopulations: grey *(64 836 animals),* bay *(33 633),* black *(9414),* chestnut *(1243),* buckskin *(433),* roan *(107),* isabella *(57) and* white *(37). Contributions to the total genetic diversity were first assessed in the existing subpopulations and later compared with two scenarios with equal subpopulation size, one with the mean population size (13 710) and another with a low population size (100). Ancestor analysis revealed a very similar origin for the different groups, except for six ancestors that were only present in one of the groups likely to be responsible for the corresponding colour. The coancestry matrix showed a close genetic relationship between the* bay *and* chestnut *subpopulations. Before adjustment, Nei's minimum distance showed a lack of differentiation among subpopulations (particularly among the* black, chestnut *and* bay *subpopulations) except for* isabella *and* white *individuals, whereas after adjustment,* white, roan *and* grey *individuals appeared less differentiated. Standardization showed that balancing coat colours would contribute preserving the genetic diversity of the breed. The global genetic diversity increased by 12.5% when the subpopulations were size standardized, showing that a progressive increase in minority coats would be profitable for the genetic diversity of this breed. The methodology developed could be useful for the study of the genetic structure of subpopulations with unbalanced sizes and to predict their genetic importance in terms of their contribution to genetic variability.*

Keywords: pedigree information, genetic variability, subpopulation size, spanish purebred horse, coat colour

Implications

It is important to quantify the contributions made to genetic diversity in order to implement management strategies in conservation programmes, which select those subpopulations that increase or maintain it. Developing a reliable method to account for these contributions would be of great interest to breeders. Changes in the structure of subpopulations within a non-endangered breed population could produce a decrease of genetic diversity and the appearance of the depression effects of inbreeding.

The present methodology has been applied to the Spanish Purebred horse to illustrate the example of a balanced representation of colour. The results showed that a progressive increase in minority coats would be profitable for the genetic diversity of this breed.

† E-mail: ebartolome@us.es

Introduction

It is important to quantify the contributions made to diversity in order to implement management strategies in programmes for the conservation of genetic variability. This task can be approached mainly from two types of material, pedigree and molecular information. A number of methods have been developed to take advantage of molecular information, including phylogenetic-like approaches, such as that of Weitzman (1992), Thaon d'Arnoldi *et al.* (1998) or those focusing on the maximization of gene diversity (Eding and Meuwissen, 2001; Ollivier and Foulley, 2005 and 2008) or

the average number of alleles per locus (Petit *et al*., 1998; Foulley and Ollivier, 2006). The coancestry-based method can be easily adapted to genealogical information (Caballero and Toro, 2002). This method is based on the comparison of between- and within-subpopulation mean coancestries. It is important to note that genealogical mean coancestries are influenced by subpopulation size. Self-coancestries have a major influence on mean coancestry values when subpopulation sizes are small. When the subpopulation sizes are not proportional to the existing metapopulation size, the results may be biased.

In horses, coat colours have been studied as a marker of population structure (Druml *et al*., 2009). Different breeding objectives associated to coat colour preferences may give some degree of population subdivision (Stachurska and Brodacki, 2000; Stachurska *et al*., 2005; Glazewska and Gralak, 2006).

One example of how to study the genetic importance of subpopulations can be found in the coat colours of the Spanish Purebred horse (SPB). From 1970 onwards, the *chestnut* coat was the only colour not allowed in the SPB studbook, whereas *grey* and *bay* coats were the most popular. However, some diluted *chestnut* coats (Bowling, 2000; Mariat *et al*., 2003) like *isabella* and *cremello* were permitted, although not encouraged. From the late 1990s onwards, however, breeders' preferences changed and the SPB became a multi-coloured breed, with an increasing number of individuals with previously less frequent coats, particularly *black*. The SPB population has been characterized as a genetically homogeneous population because of the small number of founder individuals influencing its pedigree (Valera *et al*., 2005). If individuals with minority coats such as *roan*, *white* or the diluted ones (*buckskin*, *palomino* or *isabella*) are a result of particular founder contributions, the increase in frequency of these coats in the population may balance the founder contributions at the breed level and, therefore, maintain the genetic stock.

The aim of this work was to assess the genetic contribution of each subpopulation using pedigree information, applying a modification of the methodology proposed by Caballero and Toro (2002), allowing for adjustments for size the within-subpopulation mean coancestries. It was carried out by measuring the existing genetic composition of different SPB subpopulations defined according to coat colour. A hypothetical scenario in which coat frequencies are balanced at the population level was also discussed.

Material and methods

Data

The data registered in the stud-book of the Spanish Purebred Horse from its foundation in 1970 up until 2006 were analysed. The data included 166 264 (80 355 male and 85 909 female) registered horses.

Individuals born in the past 11 years, 1996 to 2006, were selected as a reference population near the last generation interval in this breed (Valera *et al*., 2005). This reference population included 66% of the overall population (109 760 animals, 53 771 males and 55 989 females).

Individuals in the reference population were grouped in eight subpopulations according to coat colour: grey, with 64 836 (59.07%) in the reference population; bay, with 33 633 (30.64%) individuals; black, with 9414 (8.58%); chestnut, with 1243 (1.13%); buckskin, with 433 (0.39%); roan, with 107 (0.10%); white, with 37 (0.03%) and Isabella, with 57 (0.05%) individuals.

Note that this data set must not be understood as a structured population in the strict sense – however, it can be useful to illustrate the methodology, we are consciously turning a blind eye to the known fact that coat colours are defined by a few genes of major effect (Lamoreux *et al*., 2001; Gutiérrez-Gil *et al*., 2007). In practice, breeders tend to mate individuals within those of the same coat colour group, which implies a stronger relation than just a random one. What makes this population particularly interesting and useful is the wide variety in the sizes of the groups of individuals, which makes applying the methodology here particularly suitable. It should be noted that the methodology used is ultimately based on the differences found 'between' and 'within' subpopulation coancestry coefficients. It is, therefore, applicable whenever these differences exist. Nevertheless, we have used the terminology for subpopulations throughout the text, as the methodology is even more useful in real subdivision scenarios. Thus, in the whole pedigree, 92% of the animals registered in the SPB studbook had either *grey* (107 145 animals, 64%) or *bay* (46 321 animals, 28%) coat colours. The remaining 8% had minority coat colours, with roughly 6% *black* (10 481 animals) coated and nearly 1% *chestnut* (1593 animals). Coat colours such as *buckskin* (468), *roan* (130), *white* (65) or *isabella* (61) accounted for the remaining 1% of the horses. To ensure the existence of some degree of subdivision caused by colours and that, as a consequence, the results were not of stochastic nature, an analysis was carried out, randomly assigning coats to individuals and keeping the sizes of subpopulations. This analysis revealed that coancestry within subgroups decreased by 7.6%.

Pedigree analyses

Given that coancestry depends on the pedigree depth, the number of equivalence to discrete generations (t) for each individual in a pedigree was computed as the sum of $(\frac{1}{2})^n$, where n is the number of generations separating the individual from each known ancestor (Boichard *et al*., 1997).

The genetic variability was characterized at the subpopulation level by computing the total effective number of founders (f_e) and ancestors (f_a; Boichard *et al*., 1997). Parameter f_e is the number of founders that, contributing in a balanced way, would explain the wide genetic diversity of the subpopulation. Parameter f_a refers to those individuals, whose contribution to the reference subpopulation is higher than that of their ascendants, thus making allowances for bottlenecks in the pedigree.

The number of founder genome equivalents (f_g), defined as the number of founders that would be expected to produce the same genetic diversity as in the population under study if the

founders were equally represented and no loss of alleles occurred (Ballou and Lacy, 1995), was also computed for each subpopulation as the inverse of twice the average coancestry of the individuals included in the reference subpopulation. Parameter f_g allows us to identify whether coat colour subpopulations have similar origins or not, thus complementing the information provided by the analysis of founders and ancestors.

The genetic contributions of subpopulations were assessed following Caballero and Toro (2002). The average coancestry (Malécot, 1948) ($\bar{f}$) over an entire metapopulation of N_T individuals, consisting of n subpopulations, subpopulation i with N_i breeding individuals, is:

$$\bar{f} = \frac{\sum_{i,j=1}^{n} f_{ij} N_i N_j}{N_T^2} = \frac{\sum_{i=1}^{n} f_{ii} N_i}{N_T} - \overline{\mathbf{D}} = \sum_{i=1}^{n} \frac{N_i}{N_T} \left[f_{ii} - \frac{\sum_{j=1}^{n} \mathbf{D}_{ij} N_j}{N_T} \right],$$

where f_{ij} is the average pairwise coancestry between individuals of subpopulations i and j, including all $N_i \times N_j$ pairs; f_{ii} is the average pairwise coancestry within subpopulation i and D_{ij} is Nei's minimum genetic distance (Nei, 1989) between subpopulations i and j computed as $\mathbf{D}_{ij} = [(f_{ii} + f_{jj})/2] - f_{ij}$. From the formula above, it can be noted that $\bar{f}$ is dependent on the within-subpopulation coancestry (first term in brackets) and the average distance among subpopulations (second term in brackets). A proportional contribution of each subpopulation to the global coancestry was obtained following Caballero and Toro (2002) as the average coancestry of the subpopulation minus its average distance from all the others.

Adjusting coancestry for subpopulation size

The within-subpopulation coancestry is affected by the subpopulation size as self-coancestries will carry more weight when the population is small. This is because the average self-coancestry of a subpopulation i is $s_i = \frac{1}{2}(1 + F_i)$ where F_i is the average inbreeding of the subpopulation. Thus, in the extreme case of a group of non-inbred and non-related individuals, the off-diagonal will be null and only N_i self-coancestries with value $\frac{1}{2}$ will be included in the mean, which is finally equal to $f_{ii} = \frac{N_i \frac{1}{2}}{N_i^2} = \frac{1}{2N_i}$. Obviously, as the subpopulation size decreases, the within-subpopulation coancestry increases. Note that, if the relative representation of individuals in the file does not match the actual population, the conclusions can be misleading.

Assuming that both the mean inbreeding value of individuals and the mean pairwise coancestry between different individuals in a subpopulation are constant, the within-subpopulation coancestry can be adjusted to extrapolate the desired size of M as:

$$f_{ii}^{M} = f_{ii}^{N_i} - \frac{s_i}{N_i} + \frac{s_i}{M} = f_{ii}^{N_i} + \frac{(N_i - M)s_i}{MN_i}$$

where s_i is the mean self-coancestry in the subpopulation, and $f_{ii}^{N_i}$ and f_{ii}^{M} the respective within-subpopulation coancestry means for the original sample size and the desired sample size of M.

Contribution analyses

The genetic contribution of each subpopulation to the total diversity was assessed as the loss or gain of genetic diversity in the whole population after removal of this subpopulation from the data set (Caballero and Toro, 2002). After quantifying the genetic contributions of each subpopulation with their original sizes, the analyses were recomputed after standardizing subpopulation sizes to: (a) 100 animals (simulating an endangered breed; FAO, 1998); and (b) 13 710 individuals, which is the mean subpopulation size in our data set. This will allow us to infer the importance of the subpopulations defined using coat colour in hypothetical scenarios in which coat colour frequencies are balanced.

All the analyses were performed using the program ENDOG (current version v4.8) (Gutiérrez and Goyache, 2005).

Results

Table 1 shows the contributions of the main ancestors to the reference subpopulations.

The six most important ancestors are common to all the subpopulations and explain approximately 50% of each subpopulation. There are also six ancestors that only contribute, minimally, to one subpopulation, with a maximum of 2.95%, in which the *grey* coat can be seen. Most of the contribution (approximately 25%) is made by only two ancestors, which are *grey* and *bay*, respectively. The *white* subpopulation has contributions from two ancestors that do not contribute to the other subpopulations, whereas *buckskin*, *black* and *roan* share all their ancestors with the other subpopulations.

In addition to the main ancestors, the entire subpopulations shared a similar effective number of ancestors. *Chestnut* and *bay* subpopulations showed the greatest value (20), whereas *buckskin*, *roan* and *grey* subpopulations showed the smallest (16). The concentration of the origin of genes was also assessed by calculating the founder genome equivalents (Ballou and Lacy, 1995; Table 1) before and after subpopulation size adjustment. The results were similar among subpopulations, with the *bay* coat colour having the highest number of founders both in standardized and non-standardized populations (higher than eight). The *buckskin* subpopulation had the lowest f_g values (always lower than 6.5). Correlation between the f_g values obtained after subpopulation size adjustment was 1.0 and correlation between the f_g values obtained after adjustment and the original ones was 0.86. No pattern was found between the size of the subpopulation and the founder genome equivalents, and the standardization did not substantially modify the conclusions extracted from this parameter.

Between-SPB subpopulations coancestries are illustrated in Figure 1. Unlike the between-subpopulations coancestry matrix (which is not affected by the self-coancestry values), the distances matrices changed with size standardization to 100 or 13 710 individuals. The coancestry matrix showed

Table 1 *List of the 10 ancestors (in decreasing order) contributing the most (in percentage) to each Spanish Purebred horse subpopulation defined by coat colour. Genealogical parameters for each subpopulation are also given*

Id	Sex	Year of birth	White	Isabella	Roan	Buckskin	Chestnut	Black	Bay	Grey
916	M	1927	13.52	11.99	12.20	15.16	13.88	12.79	14.72	16.04
2179	F	1923	11.60	10.93	10.86	12.83	12.38	11.31	12.77	13.11
5124	M	1949	8.19	8.52	8.73	5.66	5.94	8.47	6.45	4.86
893	M	1921	3.91	4.17	3.93	6.58	6.54	5.27	5.68	4.86
947	F	1913	5.79	5.82	5.49	6.38	6.47	5.57	5.53	6.21
2735	M	1933	6.10	5.62	5.38	4.85	4.89	3.93	4.22	6.54
693	F	1919	–	–	2.65	3.50	3.40	2.73	2.94	3.68
2143	M	1914	2.85	3.21	3.20	2.75	–	3.22	–	2.68
190	F	1908	3.03	–	–	2.98	3.43	–	2.99	–
4440	M	1944	–	3.18	3.05	–	–	3.05	–	–
1382	F	1919	–	–	–	3.40	–	–	2.49	3.32
40074	M	1992	4.52	–	–	–	–	–	7.75	–
3247	M	1922	–	2.78	–	–	–	2.56	–	–
5034[a]	F	1948	–	–	–	–	2.92	–	–	–
712[a]	F	1909	–	–	–	–	2.72	–	–	–
3154[a]	F	1927	–	2.74	–	–	–	–	–	–
51218[a]	M	1995	2.72	–	–	–	–	–	–	–
349[a]	M	1907	–	–	–	–	–	–	–	2.95
3834[a]	M	1934	–	–	2.90	–	–	–	–	–
f_a			**18**	**20**	**20**	**16**	**18**	**19**	**16**	**16**
f_g		NS	6.40	6.43	6.81	6.32	8.60	8.57	8.84	7.00
		100	7.19	6.76	6.78	6.03	7.97	7.90	8.13	6.55
		13 710	7.74	7.24	7.27	6.41	8.66	8.57	8.84	7.00
E_{nf}			40	44	38	39	49	48	50	37
E_{na}			18	18	16	16	20	19	20	16

f_a = Effective number of ancestors; f_g = Founder genome equivalent obtained with and without size standardization of the reference subpopulation; E_{nf} = Effective number of founders for each subpopulation; E_{na} = Effective number of ancestors for each subpopulation.
NS = Non-standardized.
[a]Ancestors that contribute to only one subpopulation.

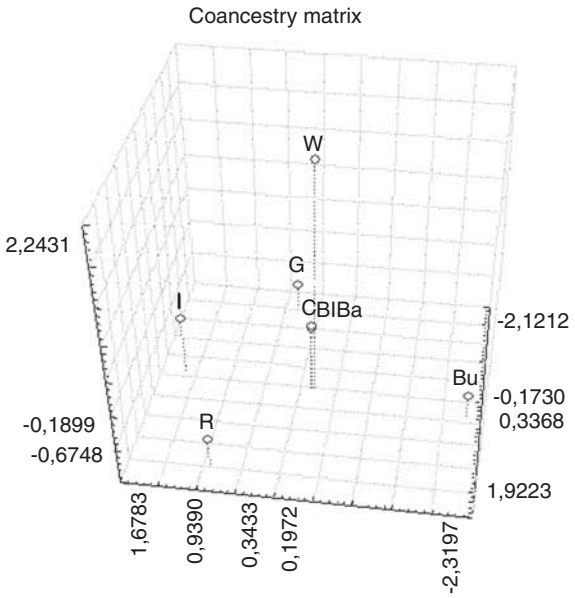

Figure 1 Three-dimensional scaling plot showing differences summarizing the information provided by coancestry matrix. **W** = White; **I** = Isabella; **R** = Roan; **Bu** = Buckskin; **C** = Chestnut; **Bl** = Black; **Ba** = Bay; **G** = Grey.

a close genetic relationship between the *bay* and *chestnut* subpopulations, on the one hand, and to a lesser extent, between the *roan* and *white* subpopulations, on the other. Before adjustment, the Nei's minimum distance showed a lack of differentiation among subpopulations (particularly among the *black*, *chestnut* and *bay* subpopulations) except for *isabella* and *white* individuals, whereas after adjustment, *white*, *roan* and *grey* individuals appeared less differentiated.

The results in Table 4 showed that, before standardization, the proportional contribution of each subpopulation to the global metapopulation was low in all cases, and was directly related to the size of the subpopulation. The *grey* subpopulation contributes the most (4.137%) to the overall coancestry when the population size is not standardized. However, these values decreased dramatically when the population size was standardized, either to a smaller size or to the mean. The relative importance of minority coats such as *buckskin*, *isabella* or *roan* increased with standardization. The correlation between internal coancestries and proportional contributions after standardization was 1.0. However, the correlations between unadjusted and adjusted values were 0.859 and 0.566, for internal coancestries and proportional contributions, respectively, showing

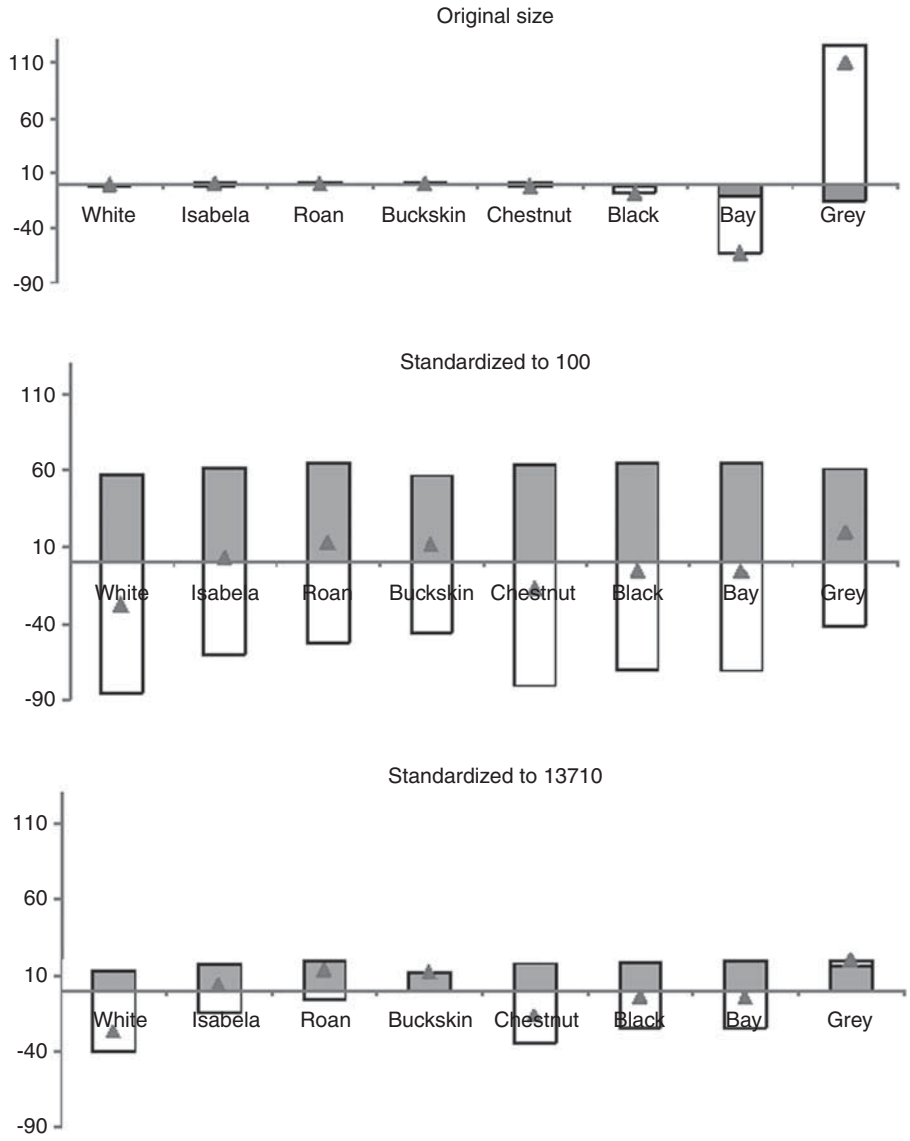

Figure 2 Bar graph indicating contributions to internal diversity (in white), between-subpopulation differentiation (in grey) and total contribution (in triangle).

that size standardization has a greater influence on subpopulation differentiations than coancestries.

The contributions of the defined subpopulations to the overall genetic diversity before and after adjustment for population size are illustrated in Figure 2. Negative values indicate that a loss of diversity will occur when the subpopulation is removed and, thus, it would be preferred for conservation purposes. When the original subpopulation sizes were considered, only *bay* and *grey* subpopulations made noticeable contributions to diversity. However, after standardization: (a) the contributions of both the *grey* and *bay* subpopulations were clarified, although both the direction and importance continued to be basically the same; (b) the *white* and *chestnut* subpopulations increased their importance so as to maintain the overall genetic diversity in SPB; and (c) the *isabella* and *roan* subpopulations did not play an important role in maintaining the overall diversity. The global genetic diversity increased when subpopulation sizes were standardized either to 100 or to the reference subpopulation mean. The results of quantification of the genetic contributions to the whole population showed that some minority coats, such as *buckskin* and *isabella,* which if excluded, made the genetic diversity decrease in the whole SPB population, changed their sign from negative to positive after standardization. However, the *white* coat increased in importance with standardized sizes as the genetic diversity changed from very low loss (0.18%) to losses of 27.85% for 100 individuals and 26.89% for 13 710 animals.

Discussion

Quantifying the genetic importance of populations or subpopulations within a metapopulation is an important research topic that has been dealt with at length in the past. This task

can be approached mainly from two types of material: pedigree and molecular information. One of the most commonly used methodologies, that of Caballero and Toro (2002), can be used from both types of information by means of the manipulation of the coancestry matrix, regardless of the source of its molecular or genealogical coancestry. When using this methodology, the comparison of between- and within-coancestry coefficients is the key information used to establish conclusions. The greater the differences in these coancestries, the greater the distances between the subpopulations and thus, the greater the subdivision. However, within the subpopulation, mean coancestry is heavily influenced by the size of the subpopulation, particularly if the different subpopulations are highly unbalanced, and this can lead to biased conclusions if such an unbalance is not a reflection of the real situation.

We have illustrated here how derived conclusions can be extremely different when considering different subpopulation sizes under the same mean pairwise coancestry and individual inbreeding coefficients. To do this, we have first discussed the genetic structure of the SPB breed regarding coat colours and then, studied the effect of balancing their representation on the SPB breed, as this representation can be defined simply by the evolution of market, tastes and fashion. Obviously, the separation produced as a consequence of the coat colour differentiation is slight and no great losses of genetic variability will be expected as a consequence of removing one of the colour groups. However, the imbalance in the subpopulation sizes has been useful to illustrate the methodology.

Although coat colour is a selection criteria that has been demonstrated not to influence horse performance (Stachurska *et al*., 2007), it is well known that market preferences exert a great influence on animal breeding (Arnason and Van Vleck, 2000; Druml *et al*., 2009). Buyers' preferences have been studied in other horse breeds such as Hanoverian, where rare coats are seen as favourites, because of the sparse offer on the market (Icken *et al*., 2007).

Nowadays, in spite of the substantially different census, it is difficult to distinguish the different coat colour subpopulations from their genetic composition; when looking at their main influencing ancestors, they often share a common origin (Table 1). In addition, the conserved genetic stock is extremely similar, as shown by the effective number of founders (37 to 50), ancestors (16 to 20) and founder genome equivalents (6.3 to 8.8, Table 1). Valera *et al*. (2005) found two major ancestors with a contribution of 15.8% and 12.6% in the SPB population, which agrees with the results about ancestors shown in Table 1. The Lipizzan horse, closely related to the SPB, had only one ancestor contributing most of the genetic variability of the breed, with 10.74% (Zechner *et al*., 2002). To summarize this information, no relationship seems to exist between the size of the subpopulation and the parameters measuring their genetic variability. Thus, *chestnut*, *bay* and *black* subpopulations, of intermediate size, showed the highest genetic diversity, whereas *buckskin*, *roan*, of smaller size, and *grey* subpopulations, the biggest, showed the lowest genetic diversity, although all of them seem to share the same genetic origin. Regarding a possible different genetic origin for these coats, in order to establish a new colour in this breed, it is essential to cross a SPB horse with an outside breed with the new coat colour (and, therefore, the new gene), but recessive alleles related to some coats could be already present at a low frequency in the population. As commonly reported in other horse breeds with long or short pedigrees, a small number of individuals may have a great influence on the breed (Zechner *et al*., 2002; Glazewska and Jezierski, 2004). In our study, this fact could have been highlighted by coat colours such as *white*, *chestnut* or *isabella*, and also with the majority *grey* colour. These coats showed ancestors that contributed to only one of these subpopulations, but with a maximum contribution of only 2.95% for the last coat. The coat colours of these marginal ancestors were among the most frequent in the SPB population and they were also different from those, which define their subpopulation, which makes them unlikely to be responsible for their introduction. Therefore, in our study, it is not clear that uncommon colour genes necessarily came from foreign breeds – on the contrary, it seems that the alleles concerned were already there. SPB coat colour variety (Abad, 2006; Agüera, 2008) was dramatically reduced because of the absorption of the population by the mainly grey Carthusian strain (Valera *et al*., 2005). The close relation among coat groups suggests that recently some breeders have been trying to produce previously unpopular colours again. This will probably lead, in some genes, to a kind of endogamy justifying the present approach by the island model (Wright, 1931).

Nei's minimum genetic distances and coancestry values are the standard instruments used to investigate population subdivision and to divide genetic variation into between- and within-population components (Eding and Bennewitz, 2007); they may also be used as indicators for the relative importance of a given cluster within a population, as subdivision is one of the major factors leading to greater increases in inbreeding than those expected according to population size (Fernández *et al*., 2008).Our results showed that there is no well differentiated subpopulation based on coat colour in this breed. This reinforces the idea that the nature of these subpopulations, as coat colours are closely related according to their genetic basis, would justify the small genetic distances shown between them. With regard to Nei's minimum distances, the smaller the related subpopulations are, the bigger this value, and the higher their coancestry values (Table 2).

One way to measure the influence of one subpopulation over the others in a given population is to ascertain the genetic contribution to diversity if one or several of these subpopulations (coat colours, in this study) are removed from the whole population (Caballero and Toro, 2002). Given that, this methodology is based in the differences found in 'between' and 'within' subpopulation coancestry coefficients, and the magnitude of the loss or gain in genetic variability of each group will depend on these differences. The coat-based subpopulations analysed here have many genetic connections, and therefore the results show that no great changes are found when one of them is removed. In spite of the fact that all the subpopulations seem to be very

Table 2 *Average coancestry among subpopulations (%, diagonal and above), Nei's minimum distance between subpopulations (%, below diagonal), and proportional contribution of each subpopulation to the global coancestry (last column) in the reference subpopulation*

	White	Isabella	Roan	Buckskin	Chestnut	Black	Bay	Grey	Proportional contribution
White	**7.81**	6.48	6.65	6.87	5.96	6.02	5.92	6.74	0.002
Isabella	1.314	**7.78**	6.39	7.19	5.86	5.91	5.83	6.47	0.004
Roan	0.919	1.166	**7.34**	6.76	6.12	6.13	6.04	7.02	0.007
Buckskin	0.996	0.657	0.864	**7.91**	6.01	6.07	5.97	6.87	0.028
Chestnut	0.850	0.936	0.459	0.850	**5.81**	5.70	5.67	6.15	0.063
Black	0.805	0.897	0.462	0.807	0.122	**5.84**	5.68	6.16	0.482
Bay	0.807	0.888	0.455	0.816	0.060	0.067	**5.65**	6.08	1.673
Grey	0.736	0.992	0.223	0.659	0.329	0.324	0.318	**7.14**	4.137

Table 3 *Nei's minimum distance between subpopulations (%) after standardization to common size of 100 individuals (below diagonal) and 13 710 individuals (above diagonal)*

	White	Isabella	Roan	Buckskin	Chestnut	Black	Bay	Grey
White		0.203	0.130	0.266	0.158	0.131	0.134	0.063
Isabella	0.700		0.497	0.165	0.481	0.460	0.452	0.556
Roan	0.509	0.994		0.577	0.209	0.229	0.225	−0.007
Buckskin	0.762	0.661	1.073		0.776	0.751	0.761	0.605
Chestnut	0.654	0.977	0.705	1.272		0.103	0.043	0.312
Black	0.627	0.956	0.725	1.247	0.600		0.067	0.324
Bay	0.630	0.948	0.721	1.257	0.539	0.564		0.321
Grey	0.560	1.053	0.489	1.101	0.809	0.821	0.817	

similar, the analyses carried out without standardization revealed that the main coat in the SPB (*grey*) was not necessary for the conservation of the population as its proportional contribution to global coancestry was 2000× greater than that for the *white* coat (Table 2), and removing it would increase the global diversity by 1%, whereas all the others except *roan* would lead to a reduction in the global genetic diversity (Table 5, column 3). Oldenbroek (2007) stated that the genetic diversity within a farm animal species is the resource by which required changes in the phenotypic characteristics of a population are made and, obviously, the more numerous ones contribute less to global diversity.

In order to highlight the genetic structure of the SPB population by means of coat colour subpopulations with size standardizations, parameters such as coancestries, Nei's minimum distances and F-statistics were also calculated after size standardization (Table 3).

The results showed that the differences in coat colour contributions increased the relative importance of minority coats such as *buckskin*, *isabella* or *roan* when the subpopulations were standardized to a small size (100 animals), or to the mean size (13 710 animals).

The conclusions we can draw from the proportional contribution to the coancestry are still more dramatic, balancing the results and changing the order of importance of the colours (Table 4): the *grey* coat is still the least important when the effect of size has been removed. The most important differences between subpopulations seem to be largely because of the differences in population sizes, since unbalanced population sizes seem to provide the major source of changes in the quantification of diversity from pedigrees. The adjustment for population sizes means that both the number of highly related SPB individuals eliminated from the data set and the number of lesser related individuals remaining is lower, thus lessening the impact on the overall diversity of the analysed population. In other words, the *grey* coat is the easiest to remove, but mainly because of its enormous relative size.

Table 4 *Average coancestry within subpopulations and proportional contribution to the global coancestry of all reference subpopulations with their original size (NS) and after standardization to 100 and 13 710 individuals*

	Internal coancestry			Proportional contribution		
	NS	100	13 710	NS	100	13 710
White	7.81	6.96	6.46	0.002	0.804	0.859
Isabella	7.78	7.40	6.90	0.004	0.621	0.675
Roan	7.34	7.37	6.87	0.007	0.643	0.697
Buckskin	7.91	8.30	7.80	0.028	0.640	0.694
Chestnut	5.81	6.27	5.77	0.063	0.874	0.928
Black	5.84	6.33	5.84	0.482	0.836	0.890
Bay	5.65	6.15	5.66	1.673	0.874	0.928
Grey	7.14	7.64	7.14	4.137	0.907	0.961

NS = non-standardized.

Taking into account the analyses after standardization as well, Table 5 shows that the *white* coat should definitely not be eliminated from the population, as this would generate

Table 5 *Contribution to genetic diversity (loss(−)/gain(+)) in % and ×100) in the reference SPB population when each subpopulation (coat) was removed, with their original size and standardized for 100 and 13 710 individuals*

Subpopulation	GD_T[a]	% Loss/gain GD[a]	GD_T[b]	% Loss/gain GD[b]	GD_T[c]	% Loss/gain GD[c]
White	0.939	−0.12–0.06 = −0.18	0.944	−85.48+57.63 = −27.85	0.945	−39.48+12.58 = −26.89
Isabella	0.939	+0.03–0.06 = −0.03	0.947	−60.09+61.85 = +1.76	0.948	−14.10+16.80 = +2.70
Roan	0.939	+0.05–0.02 = +0.04	0.948	−52.40+64.71 = +12.31	0.949	−6.42+19.65 = +13.23
Buckskin	0.939	+0.28–0.29 = −0.01	0.948	−46.47+56.82 = +10.35	0.949	−0.49+11.77 = +11.28
Chestnut	0.938	−2.21–0.36 = −2.58	0.945	−80.18+63.17 = −17.00	0.946	−34.17+18.12 = −16.05
Black	0.938	−6.43–1.21 = −7.64	0.946	−69.93+64.06 = −5.87	0.947	−23.94+19.01 = −4.93
Bay	0.933	−51.28–11.02 = −62.30	0.946	−70.39+64.45 = −5.94	0.947	−24.40+19.40 = −5.00
Grey	0.949	+125.03–14.91 = +110.12	0.949	−42.31+61.10 = +18.79	0.949	+3.67+16.05 = +19.72

SPB = Spanish Purebred horse.
GD_T = Total genetic diversity.
[a]Without standardization.
[b]Subpopulations standardized to 100 individuals.
[c]Subpopulations standardized to 13 710 individuals.

the greatest losses in genetic diversity when the subpopulation size was standardized, from −0.18% before standardization, either to 100 individuals (−27.85) or to the mean (−26.89).

Coats such as *buckskin*, *isabella* or *roan* would produce a gain in diversity if they were removed when standardized, which is not always the case before standardization (Figure 2).

Total genetic diversity (GD_T) is defined as a set of differences between species, breeds within species, and individuals within breeds expressed as a consequence of differences in their DNA (Eding and Bennewitz, 2007). It shows the global diversity when each of the subpopulations are removed, because the sustainable management of genetic resources is concerned with managing the diversity that is present today (Woolliams and Toro, 2007). The results showed that this parameter increased by 12.5% when the subpopulations were size standardized. These results show that a progressive increase in minority coats would be profitable for the genetic diversity of this breed.

As expected, the number of founder genome equivalents (f_g) decreased (Table 1) because of the size standardization when the new size was lower than the original, as a consequence of the greater weight of self-coancestries in the mean within-subpopulation coancestry. This was the case of the majority coats (*grey*, *black*, *bay* and *chestnut*). In contrast, this parameter tended to increase in minority coats (*roan*, *buckskin*, *isabella*). In addition, as the weight of self-coancestries is the same in standardized analyses, this parameter allows us to compare the genetic diversity of subpopulations based on the remaining original genetic stock. Therefore, the *grey* subpopulation showed the second lowest values after *buckskin,* whereas *bay*, *black* and *chestnut* were those with the highest values.

We have shown that the conclusions obtained from mean subpopulation coancestries are highly dependent on the unbalanced subpopulation size, as a consequence of the weight of self-coancestries. If this information reflects the real situation, the conclusions would be fair, that is, a group of individuals would be genetically important either because they are not closely related or because they are scarce, or both. If this is not the case, we have shown how the mean coancestries can be fitted to a desired size. This methodology can also be useful when predicting the genetic importance of the subpopulations if their relative size evolves, as we have done here to illustrate the example of a balanced representation of colour coats in the SPB population. Furthermore, this methodology could be helpful in the management of other animal species using conservation programmes with subdivided populations, as it could quantify the contribution of these subpopulations to genetic diversity in the whole population, using pedigree information. Zoo animals, for example, present common captive-breeding programmes for conservation that seek to minimize the harmful genetic changes potentially arising from loss of genetic diversity, inbreeding depression and the accumulation of new, mildly deleterious mutations (Ford, 2002). However, using different management techniques and allowing for mating between animals located in different zoos could lead to the subdivision of the population (Wang and Caballero, 1999). Quantifying the contributions to genetic diversity of these subpopulations could be carried out by management strategies in the conservation programmes of these species.

Acknowledgements

The authors wish to thank the National Association of Spanish Purebred Horse Breeders (ANCCE) for providing the genealogical database of this breed for this study.

References

Abad M 2006. El caballo en la historia de España. Eds León University. León, Spain.

Agüera E 2008. Córdoba, caballos y dehesas. Ed. Almuzara. Spain.

Arnason T and Van Vleck LD 2000. Genetic improvement of the horse. In The genetics of the horse (ed. A Bowling and A Ruvinsky), pp. 473–498. CABI Publishing, Wallingford, UK.

Ballou JD and Lacy RC 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 (ed. JD Ballou, M Gilpin and TJ Foose), pp. 76–111. Columbia University Press, New York, USA.

Boichard D, Maignel L and Verrier E 1997. The value of using probabilities of gene origin to measure genetic variability in a population. Genetics Selection Evolution 29, 5–23.

Bowling AT 2000. Genetics of colour variation. In The genetics of the horse (ed. A Bowling and A Ruvinsky), pp. 53–70. CABI Publishing, Wallingford, UK.

Caballero A and Toro MA 2002. Analysis of genetic diversity for the management of conserved subdivided populations. Conservation Genetics 3, 289–299.

Druml T, Baumung R and Sölkner J 2009. Pedigree analysis in the Austrian Noriker draught horse: genetic diversity and the impact of breeding for coat colour on population structure. Journal of Animal Breeding and Genetics 126, 348–356.

Eding H and Bennewitz J 2007. Measuring genetic diversity in farm animals. In Utilization and conservation of farm animal genetic resources (ed. K Oldenbroek), pp. 103–130. Wageningen Academic Publishers, Wageningen, The Netherlands.

Eding H and Meuwissen THE 2001. Marker-based estimates of between and within population kinships for the conservation of genetic diversity. Journal of Animal Breeding and Genetics 118, 141–159.

FAO 1998. Secondary guidelines for the national farm animal genetic resources management plans: management of small populations at risk. FAO, Rome, Italy.

Fernández J, Toro MA and Caballero A 2008. Management of subdivided populations in conservation programs: development of a novel dynamic system. Genetics 179, 683–692.

Ford MJ 2002. Selection in captivity during supportive breeding may reduce fitness in the wild. Conservation Biology 16, 815–825.

Foulley J-L and Ollivier L 2006. Estimating allelic richness and its diversity. Livestock Science 101, 150–158.

Glazewska I and Gralak B 2006. Balancing selection in Polish Arabian horses. Livestock Science 105, 272–276.

Glazewska I and Jezierski T 2004. Pedigree analysis of Polish Arabian horses based on founder contributions. Livestock Production Science 90, 293–298.

Gutiérrez JP and Goyache F 2005. A note on ENDOG: a computer program for analysing pedigree information. Journal of Animal Breeding and Genetics 122, 357–360.

Gutiérrez-Gil B, Wiener P and Williams JL 2007. Genetic effects on coat colour in cattle: dilution of eumelanin and phaeomelanin pigments in an F2-Backcross Charolais × Holstein population. BMC Genetics 8, 1–12.

Icken W, Bennewitz J and Kalm E 2007. Analysis of auction data for horses and influence factors. Züchtungskunde 79, 111–118.

Lamoreux ML, Wakamatsu K and Ito S 2001. Interaction of major coat color gene functions in mice as studied by chemical analysis of eumelanin and pheomelanin. Pigment Cell Research 14, 23–31.

Malécot G 1948. Les Mathématiques de l'Hérédité. Masson et Cie, Paris, France.

Mariat D, Taourit S and Guerin G 2003. A mutation in the MATP gene causes the cream coat color in the horse. Genetics Selection Evolution 35, 119–133.

Nei M 1989. Molecular evolutionary genetics. Columbia University Press, New York.

Oldenbroek K 2007. Introduction. In Utilisation and conservation of farm animal genetic resources (ed. K Oldenbroek), pp. 13–28. Wageningen Academic Publishers, Wageningen, The Netherlands.

Ollivier L and Foulley J-L 2005. Aggregate diversity: new approach combining within- and between-breed diversity. Livestock Production Science 95, 247–254.

Ollivier L and Foulley J-L 2008. Managing genetic diversity, fitness and adaptation of farm animal genetic resources. In Adaptation and fitness in animal populations, evolutionary and breeding perspectives on genetic resource management (ed. JHJ van der Werf, H-U Graser, R Frankham and C Gondro), pp. 201–227. Springer-Science and Business Media, BV, Australia.

Petit RJ, El Mousadik A and Pons O 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12, 844–855.

Stachurska A and Brodacki A 2000. Genetic structure of Malopolski horse population with respect to basic coat colours. Annals of Animal Science 27, 9–18.

Stachurska A, Brodacki A and Klimorowska A 2005. Genetic structure of Silesian horse population with regard to the coat colour (in Polish). Roczniki Naukowe PTZ 1, 63–72.

Stachurska A, Pieta M, Lojek J and Szulowska J 2007. Performance in racehorses of various colours. Livestock Science 106, 282–286.

Thaon d'Arnoldi C, Foulley J-L and Ollivier L 1998. An overview of the Weitzman approach to diversity. Genetics Selection Evolution 30, 149–161.

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

Wang J and Caballero A 1999. Developments in predicting the effective size of subdivided populations. Heredity 82, 212–226.

Weitzman ML 1992. On diversity. The Quarterly Journal of Economics 107, 363–405.

Woolliams J and Toro M 2007. Chapter 3: what is genetic diversity? In Utilisation and conservation of farm animal genetic resources (ed. K Oldenbroek), pp. 55–74. Wageningen Academic Publishers, Wageningen, The Netherlands.

Wright S 1931. Evolution in Mendelian populations. Genetics 16, 97–159.

Zechner P, Sölkner J, Bodo I, Druml T, Baumung R, Achmann R, Marti E, Habe F and Brem G 2002. Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. Livestock Production Science 77, 137–146.