

## ORIGINAL ARTICLE

# Pedigree information reveals moderate to high levels of inbreeding and a weak population structure in the endangered Catalanian donkey breed

J.P. Gutiérrez<sup>1</sup>, J. Marmi<sup>2</sup>, F. Goyache<sup>3</sup> & J. Jordana<sup>2</sup>

<sup>1</sup> Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense de Madrid, Madrid, Spain

<sup>2</sup> Unitat de Genètica i Millora Animal, Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>3</sup> Serida-Somió, Camino de los Claveles, Gijón, Asturias, Spain

## Correspondence

Jordi Jordana, Unitat de Genètica i Millora Animal, Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193-Bellaterra, Barcelona, Spain. Tel: +34-935811217; Fax: +34-935812106; E-mail: jordi.jordana@uab.es

Received: 31 January 2005;  
accepted: 3 June 2005

## Summary

The Catalanian donkey is one of the most endangered donkey breeds in the world. At present, five main subpopulations exist: AFRAC, which consists of many genetically connected Catalanian localities; Berga, which consists of a single herd located also in Catalunya but under private management; and three minor non-Catalonian subpopulations (Huesca, Sevilla and Toledo). In this study, we analysed the pedigree information of the Catalanian donkey herdbook to assess the genetic diversity and population structure of the breed. We found that the Catalanian donkey has suffered an important loss of genetic diversity and moderate to high increases of inbreeding because of the abuse of a few individuals in matings. This scenario is mainly characterized by the fact that both the effective number of founders and ancestors for the whole population was 70.6 and 27, respectively, while the equivalent number of founders was 146.5 and the number of ancestors explaining overall genetic variability was 93. In addition, only 14% of animals born between the 1960s and 1970s were significantly represented in the pedigree. Our results also show that subpopulations where breeders exchanged reproductive individuals had low levels of inbreeding and average relatedness. One subpopulation, Berga, was reproductively isolated and showed high levels of inbreeding ( $F = 7.22\%$ ), with average relatedness ( $AR = 6.61\%$ ) playing an important role in increasing the values of these coefficients in the whole pedigree. Using genealogical  $F$ -statistics we have found little evidence of population structuring ( $F_{ST} = 0.0083$ ) with major genetic differences among non-Catalonian subpopulations.

## Introduction

The Catalanian donkey is an extremely endangered breed present mainly in the Pyrenean and pre-Pyrenean areas of Catalonia (north-eastern Spain) that probably descended from the ancestral lineage belonging to the Somalian donkey (*Equus asinus somaliensis*) (Epstein 1984; Aranguren-Méndez *et al.*

2004; Beja-Pereira *et al.* 2004). The Catalanian donkey breed was close to extinction at the end of the Spanish Civil War (1936–1939). Although the population recovered between the 1940s and 1950s, it decreased alarmingly again, because of an economic crisis that affected the equine sector, and rural mechanization and depopulation. In 1978, the 'Associació del Foment de la Raça Asinina Catalana

(AFRAC)' was founded and the Catalanian donkey herdbook was reopened.

Assessment and conservation of genetic variability in minor livestock breeds, such as the donkey, is a growing worldwide concern because of the high risk of breed loss (Aranguren-Méndez *et al.* 2001). At the end of 1994, the Catalanian donkey breed Conservation Program was started with two priorities: (1) maintaining maximum genetic diversity, with (2) the minimum possible consanguinity increase per generation (Jordana & Folch 1998). However, in such a small population, matings cannot be planned using only unrelated animals, leading to an increase of overall inbreeding, the reduction of the effective population size and, consequently, the loss of genetic variability (Falconer & MacKay 1996). Hence at a demographic level, the implementation of a mating policy ensuring the equal contribution of the maximum number of animals from both sexes to the next generation, the minimization of inbred matings, and the maximization of the genetic representation of the founder individuals in the next generation was proposed (Folch & Jordana 1998).

Molecular studies allow us to conclude that the inbreeding of the Catalanian donkey breed is not negligible and is probably increasing. At the beginning of the Conservation Program, Jordana *et al.* (1999), using microsatellites and protein markers, obtained non-significantly different from zero-average  $F_{IS}$  values, but Jordana *et al.* (2001), with a higher sample of individuals genotyped for 10 microsatellites, found an average  $F_{IS}$  value of 0.154. More recently, Aranguren-Méndez *et al.* (2002) with 15 microsatellites obtained an average  $F_{IT}$  value (as an estimation of the total inbreeding of the breed) of 0.232.

Recent studies (Caballero & Toro 2000, 2002), have formalized useful tools to assess, using genealogical information, the genetic structure of a population arising from local bottlenecks or a deficient mating policy. The aim of this study was to analyse the genetic variability of the Catalanian donkey breed in terms of inbreeding and genetic representation and to ascertain the possible presence of genetic structuring in the breed. A number of genealogical parameters that were not computed before in the breed have been obtained. This analysis will test the reliability of current mating strategies to conserve the genetic variability of the breed.

## Material and methods

We analysed the pedigree information registered in the Catalanian donkey herdbook from its foundation

in the late 1970s to 31 December 2002. A total of 510 (300 females) animals were registered, of which 336 individuals (250 females) were alive. According to the geographical location, up to five subpopulations can be distinguished in the Catalanian donkey breed: AFRAC, Berga, Huesca, Sevilla and Toledo, including 229, 193, 26, 46 and 16 registered individuals respectively. The subpopulations are basically formed by small herds grouped according to their geographical proximity. The latter three subpopulations are located outside Catalonia. The AFRAC subpopulation forms most of the herds in Catalonia. Despite its location in Catalonia, the Berga subpopulation consists of a single herd under private management, thus leading it to be characterized as a different subpopulation.

The pedigree completeness level was characterized by computing:

1 *Generation lengths*. This is the average age of parents at the birth of the offspring used to replace them. We computed this for the four pathways (father-son, father-daughter, mother-son and mother-daughter) using birth dates of registered animals together with those of their fathers and mothers.

2 *Number of fully traced generations ( $g$ )*. This is computed as the value separating the offspring of the farthest generation where the  $2^g$  ancestors of the individual are known. Ancestors with no known parent were considered as founders (generation 0).

3 *Number of equivalent generations ( $g_e$ )*. This is computed as the sum of  $(1/2)^n$  where  $n$  is the number of generations separating the individual from each known ancestor.

To characterize the genetic variability of the population, we analysed the following parameters:

1 *Effective number of founders ( $f_e$ )* (James 1972; Lacy 1989). This is defined as the number of equally contributing founders that would be expected to produce the same genetic diversity as in the population under study. It is computed as:

$$f_e = \frac{1}{\sum_{k=1}^f q_k^2}$$

where  $f$  is the number of founders and  $q_k$  is the probability of gene origin of the  $k$  ancestor, that is the fraction of population originated from the  $k$  ancestor.

2 *Effective number of ancestors ( $f_a$ )*. This is the minimum number of ancestors, not necessarily founders, explaining the complete genetic diversity of a population (Boichard *et al.* 1997). Parameter  $f_a$  does not fully account for gene loss by drift from the

ancestors to a reference population but complements the information offered by the effective number of founders accounting for the losses of genetic variability produced by the unbalanced use of reproductive individuals producing bottlenecks. To compute  $f_a$ , one must consider only the marginal contribution of an ancestor, which is the contribution made by an ancestor that is not explained by other ancestors chosen before. We computed this by following the algorithm described by Boichard *et al.* (1997).

**3 Individual inbreeding coefficient ( $F$ ).** It is defined as the probability that an individual has two identical genes, by descent (Wright 1931).

**4 Effective population size ( $N_e$ ).** It is defined as the number of breeding animals that would lead to the actual increase in inbreeding if they contributed equally to the next generation.  $N_e$  was computed as  $N_e = 1/2\Delta F$  (Wright 1931) for the pedigree completeness level and for each fully traced generation level. The increase of inbreeding ( $\Delta F$ ) was calculated by means of the classical formula (Wright 1931):

$$\Delta F = \frac{F_t - F_{t-1}}{1 - F_{t-1}}$$

where  $F_t$  and  $F_{t-1}$  are the average inbreeding and heterozygosity at the  $i$ th generation. However, this usually fits poorly to real populations giving an overestimate of the actual effective population size (Goyache *et al.* 2003). In consequence, the increase in inbreeding between two generations ( $F_n - F_{n-1}$ ) was assumed to be the regression coefficient ( $b$ ) of the individual inbreeding coefficient over the number of fully traced generations and the number of equivalent generations.

**5 Average relatedness coefficient (AR) of each individual** (Goyache *et al.* 2003; Gutiérrez *et al.* 2003). This is computed as the average of the coefficients in the row corresponding to the individual in the numerator relationship matrix **A**. Thus, the AR coefficient accounts simultaneously for the co-ancestry and inbreeding coefficients. The AR coefficients can then be interpreted as the representation of the animal in the whole pedigree. Obviously, the addition of the AR scores across founder animals is equal to 1 (100% of the population).

The reference population used to compute the parameters above is the whole known pedigree (given its small size) including founders. However, to compute parameter  $f_a$ , we used, as a reference population, only the non-founder animals.

Population structure was assessed from genealogical information by means of  $F$ -statistics (Wright 1978) for each fully traced generation and each

defined subpopulation. Wright's  $F$ -statistics have been computed following Caballero & Toro (2000, 2002). These authors have formalized the pedigree tools necessary for the analysis of genetic differentiation in subdivided populations starting with the average pairwise co-ancestry coefficient ( $f_{ij}$ ) between individuals of two subpopulations,  $i$  and  $j$ , of a given metapopulation including all  $N_i \times N_j$  pairs. For a given subpopulation  $i$ , the average co-ancestry, the average self-co-ancestry of the  $N_i$  individuals and the average coefficient of inbreeding would be  $f_{ii}$ ,  $s_i$  and  $F_i = 2s_i - 1$  respectively. From these parameters and the corresponding mean values for the entire metapopulation Caballero & Toro (2000, 2002) obtained the Wright's (1978)  $F$ -statistics as

$$F_{IS} = \frac{\bar{F} - \bar{f}}{1 - \bar{f}}, \quad F_{ST} = \frac{\bar{f} - \bar{f}}{1 - \bar{f}} = \frac{\bar{D}}{1 - \bar{f}}, \quad \text{and} \quad F_{IT} = \frac{\bar{F} - \bar{f}}{1 - \bar{f}},$$

where  $\bar{f}$  and  $\bar{F}$  are the mean co-ancestry and the inbreeding coefficient for the entire metapopulation, respectively, and,  $\bar{f}$  the average co-ancestry for the subpopulation so that  $(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$ . Moreover, a generic distance between individuals  $i$  and  $j$ , here called kinship distance ( $D_k$ ), corresponding to the same subpopulation would be  $D_k = \lfloor (s_i + s_j)/2 \rfloor - f_{ij}$ . Within breeds,  $D_k$  were simply computed averaging the corresponding values for all the within- or between-subpopulation pairs of individuals.

Most parameters have been computed using the program ENDOG 3.0 (Gutiérrez & Goyache 2005). A phylogenetic tree using  $F_{ST}$  distances was drawn using MEGA 2.1 program (Kumar *et al.* 2001). When needed, statistical analysis was carried out using the SAS program (SAS/STAT 1999).

## Results

At population level, the number of registered individuals with unknown parents was 128 (41 males and 87 females). Up to 37 additional registered individuals had one parent unknown and were included in the baseline population. Thus, weighting the unknown parent as half a founder, the equivalent number of founders ( $f_{te}$ ) would increase to 146.5. The number of ancestors (founders or not) necessary to explain the genetic variability of the breed was 93. In fact, only 11 and 24 ancestors were necessary to explain 50 and 70% of the present genetic variability in the Catalanian donkey respectively. The effective number of founders ( $f_e = 70.6$ ) was half of the equivalent number of founders whereas the

effective number of ancestors ( $f_a = 27$ ) was between a third and a quarter of the number of ancestors explaining overall genetic variability. Average  $F$  and AR for the whole population were 3.36 and 3.76% respectively. Average number of fully traced generations and average generation equivalents were 1.23 and 1.96 respectively.

Most of the parameters above referred to the different subpopulations defined in the Catalanian donkey breed are given in Table 1. Within subpopulation, mean inbreeding coefficients for living animals were 11.04% for Berga, 3.1% for Sevilla, 2.1% for Toledo, 1.1% for AFRAC and 0.5% for Huesca. The AFRAC and Berga subpopulations are the most important. They summed up more than 80% of the individuals, equivalent number of founders and ancestors of the herdbook. The founders from the AFRAC subpopulation contributed to the genetic pool of the Catalanian donkey in 53.43%. However, when all the individuals were considered (regardless whether they were founders or not), the Berga subpopulation made the major genetic contribution to the gene stock (51.44%). The influence of the Toledo subpopulation on the breed was negligible both in absolute and relative terms. The Berga subpopulation had the highest inbreeding and AR values (7.22 and 6.61% respectively). This is consistent with the deeper pedigree knowledge found for this subpopulation (2.78 equivalent generations). Average inbreeding was below 1% for the AFRAC and

Huesca subpopulations and between 2 and 3% for the Sevilla and Toledo subpopulations. Sevilla showed moderate levels of genetic representation whereas AR values were lower in AFRAC, Huesca and Toledo subpopulations. The Berga and (to a lesser extent) Sevilla and AFRAC subpopulations had positive  $F_{IS}$  values showing that within these subpopulations the average  $F$  exceeded the between-individual co-ancestry. However, the same parameter for the Huesca subpopulation showed the opposite pattern ( $F_{IS} = -4.46 \times 10^{-2}$ ). This situation is also reflected by the within-subpopulation  $D_k$ . This parameter was the highest and of similar magnitude for the AFRAC and Berga subpopulations showing that the within-subpopulation self-co-ancestry exceeded the genetic similarity between individuals.

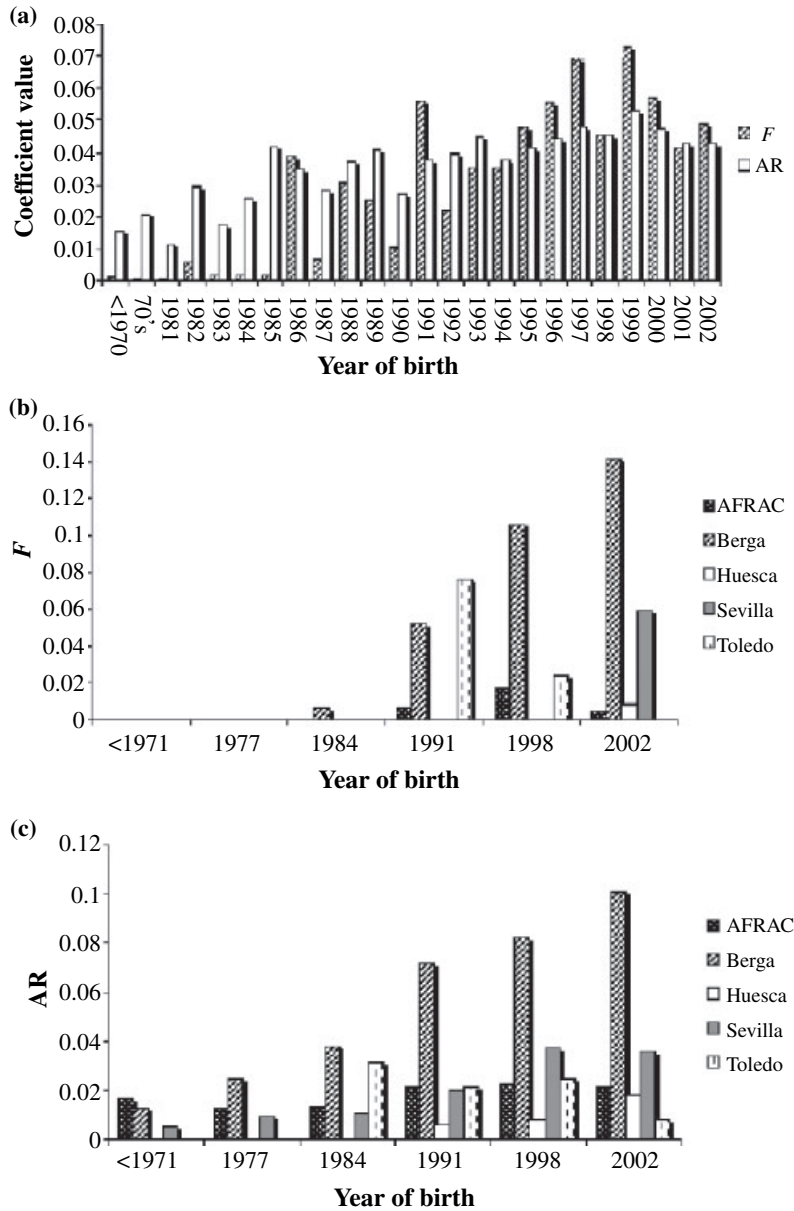
The pedigree completeness level is described in Table 2. Inbreeding and AR values were higher in males than in females and in alive than in dead animals (Table 2). Figure 1(a) shows that total  $F$  increased quickly from the second half of the 1980s to be between 4 and 7% in individuals born since 1995. Average of AR values were more stable, reaching values between 4 and 5% from 1985. In any case, average  $F$  has tended to be higher than the corresponding AR values since the second half of the 1990s. The Berga subpopulation showed the highest increases and values of  $F$  and AR whereas the same coefficients remained at very low values or constant in AFRAC (Figure 1b,c). Toledo and Sevilla showed

**Table 1** Main genealogical parameters computed for the whole population and for each of the subpopulations included in the Catalanian donkey breed herdbook

Population	Whole	AFRAC	Berga	Huesca	Sevilla	Toledo
Total number of animals	510	229	193	26	46	16
Animals with unknown parents	128	75	33	8	7	5
Equivalent number of founder animals	146.5	85.5	39.5	8	8	5.5
Genetic contribution of the founders (%)	100	53.43	34.45	4.45	6.01	1.66
Number of ancestors	93	48	28	7	6	4
Variability explained by the ancestors (%)	100	36.52	51.44	6.91	4.26	0.87
$F$ (%)	3.36	0.73	7.22	0.48	2.17	2.81
AR (%)	3.76	1.98	6.61	1.39	2.78	1.66
Fully traced generations	1.23	0.83	1.83	0.81	1.15	0.81
Equivalent generations	1.96	1.39	2.78	1.20	2.02	1.39
$F_{IS}$ (x100)	2.89	0.37	3.59	-4.46	1.32	-0.67
Within $D_k$	0.514	0.513	0.510	0.457	0.492	0.470

**Table 2** Average values of inbreeding ( $F$ ), average relatedness (AR), number of fully traced generations, number of equivalent generations ( $E_q$ ) and equivalent to discrete generations ( $T$ ) for the Catalanian donkey pedigree completeness level and for the inbred animals

	Whole pedigree	Inbred	Males	Females	Alive	Dead
n	510	137	210	300	336	174
$F$ (%)	3.36	12.54	4.05	2.88	4.71	0.59
AR (%)	3.76	8.39	4.45	3.27	4.61	2.14
Fully traced generations	1.23	2.61	1.39	1.12	1.58	0.58
Generations equivalent	1.96	3.93	2.22	1.77	2.50	0.93



**Figure 1** Average values of inbreeding (*F*) and average relatedness (AR) coefficients by year of birth of animals for total population (a) and for each subpopulation alone (b, c).

moderate increases of *F* – reaching 8% at the beginning of the 1990s and 6% at 2002 – in non-Catalonian subpopulations (Figure 1b). AR coefficients of non-Catalonian subpopulations did not reach 4% and showed little increases or decreased since the mid-1980s (Figure 1c). Effective population size computed from the average *F*-value for the whole population was 14.9. The  $N_e$  obtained from the regression coefficient of the individual inbreeding coefficient over the number of fully traced generations and the number of equivalent generations were 12.5 and 17.7 respectively.

Realized *F*-value increased quickly with pedigree knowledge (Table 3). Average *F*-value for individuals

with more than three generations in their genealogies reached 20.7%. A similar pattern was observed for AR, but it was opposite for  $N_e$ . Genealogical  $F_{IS}$  values computed by generation are basically 0 till generation 2. After this generation,  $F_{IS}$  values tended to show a clear co-ancestry excess with values of  $-5.35 \times 10^{-2}$  and  $-12.26 \times 10^{-2}$  for animals with three and more than three fully traced generations respectively. Table 3 shows differentiation between generations assessed between pairwise  $F_{ST}$  values ranging from 0.05% (between the founder generation and the first generation) to 3.05% (between the founder generation and the last generation). Generation 1 is highly differentiated from the subsequent

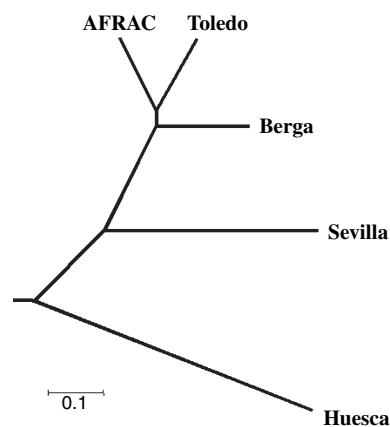
**Table 3** Variables characterizing genetic variability in the Catalanian donkey by number of fully traced generations

	Individuals		$F$ (%)	AR (%)	$N_e$	$F_{IS}$ ( $\times 100$ )	Paired $F_{ST}$ ( $\times 100$ ) distances by generation			
	N	Inbred percentage					0	1	2	3
0	165	0	0.00	0.97		-0.09				
1	160	9	0.90	2.47	55.7	-0.31	0.05			
2	105	52	4.80	5.30	12.8	1.00	0.65	1.33		
3	66	61	11.51	9.63	7.4	-5.35	0.2	1.94	0.64	
>3	15	100	20.70	11.71	5.4	-12.26	3.05	1.94	0	0

Parameters showed are: number of individuals, percent of inbred individuals, average inbreeding, average relatedness (AR), effective size and  $F_{IS}$ . Additionally, paired  $F_{ST}$  distances between generations computed from genealogical data are shown.

**Table 4** Between subpopulations paired  $F_{ST}$  values ( $\times 100$ ) computed from the genealogical information contained in the Catalanian donkey herdbook

Subpopulation	Berga	Huesca	Sevilla	Toledo
AFRAC	0.30	0.44	0.36	0.30
Berga		0.53	0.52	0.41
Huesca			1.46	2.23
Sevilla				1.46
Toledo				

**Figure 2** UPGMA tree of Catalanian donkey subpopulations based on  $F_{ST}$  distances.

generations (values from 1.33 to 1.94%) while generations 2 and 3 are virtually undifferentiated.

$F$ -statistics,  $F_{IS}$ ,  $F_{IT}$  and  $F_{ST}$ , computed from pedigree information had the following values: 2.89, 3.70 and 0.83% respectively. The between-subpopulation values of  $F_{ST}$  are given in Table 4. The AFRAC and the Berga subpopulations had the lowest differentiation (0.30%) and showed  $F_{ST}$  values with other subpopulations below the  $F_{ST}$  computed for the whole breed. The higher values of differentiation were found between the subpopulations outside

Catalonia, showing a limited gene flow between them. Figure 2 shows the phylogenetic relationships of Catalanian donkey subpopulations based on  $F_{ST}$  distances and the UPGMA method.

## Discussion

The Catalanian donkey may have suffered an important loss of founder genetic diversity, as suggested by the low value of the effective number of founders obtained ( $f_e = 70.6$ ) compared with the number of founder animals registered in the herdbook ( $f_i = 128$  and  $f_{ic} = 146.5$ ). From the information available in 1996, Folch & Jordana (1998) reported an  $f_e/f_i$  ratio equal to 0.60 that means that for every founder contributing effectively to the population genetic pool there was another for which the information was lost. In the present study, we obtained slightly lower values ( $f_e/f_i = 0.55$  and  $f_e/f_{ic} = 0.48$ ). This result was probably influenced by the registration of 43 additional founders in the herdbook since 1996. Animals with more than three fully traced generations show remarkable genetic differences regarding founders (Table 3). These results make it clear that many founder genotypes would be barely represented or erased from the population. However, when we compare the two greater subpopulations (Berga and AFRAC) we found that this picture is more complex. The AFRAC subpopulation contains the greatest number of founders and their genotypes represent a half of the Catalanian donkey founder genetic diversity (Table 1). On the contrary, founders from the Berga subpopulation represent only 30% of the founder genetic diversity of the breed. Nevertheless, in the case of the Catalanian donkey, some founders probably did not retain ancestral genotypes because, although they had no known parents, they were born after the 1980s. So, it is possible that the genetic diversity of founders

would be not exactly equivalent to the ancestral genetic diversity of the breed.

A better estimate of the losses of Catalanian donkey ancestral genetic diversity could be obtained from the genetic representation of ancestors born between the 1960s and 1970s. Eighty-eight registered animals were born between these two decades (most of them are dead) and only 12 of them appeared within the 24 ancestors that explained 70% of total genetic diversity of Catalanian donkey. Thus, only genotypes from 14% of animals born between the 1960s and 1970s were significantly represented in the pedigree explaining 38% of total genetic diversity of the breed. Nineteen animals from AFRAC were born between the 1960s and 1970s explaining 32% genetic diversity of their corresponding subpopulation. In the case of Berga, only eight individuals explained 63% of the subpopulation's genetic diversity. However, this high percentage is due to the genetic contribution of basically three animals. All these results agree with the hypothesis that ancestral genetic diversity is scarcely represented (less than a quarter of total genetic diversity) in the Catalanian donkey pedigree. On the contrary, the over-representation of only three old animals in Berga subpopulation has probably had a great effect in the loss of most of this ancestral genetic diversity of the breed.

Average inbreeding reported in the present study ( $F = 3.36\%$  for total pedigree and  $F = 4.71\%$  for living population) is lower than that previously reported ( $F = 6\%$ ) by Folch & Jordana (1998). However, this result may be doubtful. Figure 1(a) shows that total inbreeding accumulated quickly since 1986, although its value has experienced some fluctuations over time with moderate falls in 1987 and 1990 and peaks in 1991, 1997 and 1999. The negative value of  $F_{IS}$  suggests that mean co-ancestry exceeds mean  $F$  in the youngest generations (Table 3) and this allows us to predict a future increase of inbreeding when these individuals mate. The AR has exceeded 1% since the 1970s. These results prove the abusive representation of some individuals in the whole pedigree that can be highly detrimental to maintaining the genetic diversity of the breed. This is also supported by the low value of the effective number of ancestors ( $f_a = 27$ ) and that only 11 ancestors explained 50% of total genetic variability of the breed. The moderate genetic representation of inbred animals (Table 2) may also be a sign of the genetic diversity loss of the Catalanian donkey breed. Nevertheless, the picture obtained is different depending on the management strategies in Catalanian donkey

major subpopulations, AFRAC and Berga. Our results fit the management policies of these two major subpopulations. The objective of AFRAC is to preserve the genetic diversity of the breed, whereas the priority at Berga is to obtain a highly selected and morphologically homogeneous herd including closely related animals. The AFRAC subpopulation exchanges breeding animals among the different localities included in the association giving, in this case, the lowest values of  $F$  and AR of the breed (Table 1). On the contrary, the Berga subpopulation is a genetically closed herd that has only been a donor of studs, but it has not incorporated any from other subpopulations. As a result, this subpopulation reports the highest values of these coefficients. Twenty-three ancestors explained 50% of genetic diversity from AFRAC and only four in Berga. In addition, inbreeding values have been low over time in AFRAC whereas they have increased sharply in Berga since the early 1990s (Figure 1b). AR coefficient acts as a good indicator of the evolution of the future inbreeding in a population (Goyache *et al.* 2003). AR has remained quite constant over time in the AFRAC subpopulation whereas it has increased sharply in Berga since the 1970s (Figure 1c). This suggests that inbreeding may continue to increase dangerously if the breeding strategy of Berga does not change in the future. Some problems linked to inbreeding may also appear. For instance, animals from this subpopulation begin to show lower reproductive efficiency (J. Miró, personal communication).

This study proves the important role that the closed management of the Berga subpopulation has in the increase of inbreeding and the drop of genetic diversity in the Catalanian donkey breed; whereas, nowadays, most matings within AFRAC are among unrelated animals as suggested by the fact that average AR was between two and three times the average  $F$ -value (Table 1). Each non-Catalonian subpopulation represents a minor fraction of the total number of individuals of the breed. Sevilla showed a moderate increase of inbreeding in animals born at 2002 (Figure 1b). Moreover, the similar values of  $F$  and AR obtained for Sevilla, as is the case of Berga, suggest that there are matings among related animals within this subpopulation. On the contrary, in the Huesca subpopulation, unrelated matings would be more frequent. However, these minor subpopulations are vulnerable to high increments of inbreeding because of their extremely low population sizes (Table 1) if breeders do not periodically introduce individuals from other subpopula-

tions. Following matings of most genetically represented individuals in the herdbook and their descendants, we have identified four major independent lines highly represented in the current living population which may be described by the following males: Campanero, Travé, Hormigon, and Halley.

Inbreeding and AR coefficients increased with pedigree knowledge whereas effective population size ( $N_e$ ) decreased over time (Table 3). Te Braake *et al.* (1994) warned that accurate estimates of  $N_e$  cannot be obtained from short-length pedigrees, as in the Catalanian donkey in which no individual has more than five fully traced generations. However, we have obtained similar estimates of  $N_e$  from average inbreeding and from regression coefficients of individual inbreeding coefficients over the number of generations.

Weak population structuring in the Catalanian donkey and other Spanish donkey breeds was previously reported using microsatellite genetic markers (Aranguren-Méndez *et al.* 2002). We have also found weak evidence of population structuring ( $F_{ST} = 0.0083$ ) which suggests that there is a little genetic isolation among subpopulations. However, most of the paired  $F_{ST}$  values showed in the Table 4 point out the existence of non-negligible between-subpopulation differentiation (see Table 4). In consequence, population structuring can be higher than that assessed here and the  $F_{ST}$  value for the whole population can be affected by the shallowness of the Catalanian donkey pedigree. Available information on population differentiation using genealogical information is scarce. Recently, Valera *et al.* (2005) have reported a lack of differentiation between Cathusian and non-Carthusian Andalusian horses ( $F_{ST} = 0.000026$ ) using a pedigree with 14 generations traced per individual, as average. Possible explanations for the low value of whole-population  $F_{ST}$  obtained in our work may be, besides the shallowness of the available pedigree, the small sample size of the non-Catalonian subpopulations and that most individuals included within them were founders or were closely related to Catalanian animals. According to between-subpopulation  $F_{ST}$  genetic distances, AFRAC and Berga were the most closely related subpopulations (Table 4 and Figure 2). This may be due to the geographic proximity of these two subpopulations and the use of studs from Berga to mate AFRAC jennies. Slightly greater genetic distances were obtained among non-Catalonian subpopulations possibly because of geographic distance and that there is no genetic exchange among these subpopulations.

Figure 2 shows that genetic differentiation among Catalanian and non-Catalonian subpopulations was less or more important depending on the case. For instance, Toledo was closely related to the Catalanian subpopulations probably because of the fact that most of the parents of animals from Toledo are from Catalanian subpopulations. On the contrary, Huesca was the most genetically differentiated non-Catalonian subpopulation. Within Huesca, most individuals have parents from their own subpopulation. Thus, the Huesca subpopulation would be the most genetically isolated subpopulation of the breed.

In this study, we show that the use of pedigree information is relevant for the management of Catalanian donkey subpopulations. This is of major importance because, as reported by Dunner *et al.* (1998), values obtained from pedigree data may be consistent with estimates obtained from molecular data indicating the robustness of pedigree analyses and the compatibility between these two sources of information. The shallowness of our pedigree directs us to consider the present results carefully. However, we are confident on that the analysis of pedigree information can highlight possible losses of genetic diversity (mainly because of genetic drift) in early stages of a conservation programme with limited costs. According to our results, we suggest that a breeding strategy to minimize the loss of genetic variability in the Catalanian donkey breed in the future should be based on the following points:

- Ensure the equal contribution of the maximum number of animals possible, mainly those that are less represented in the pedigree – showing the lowest values of AR – leaving offspring to the next generation. The use of some programs, such as the GENCONT software (Meuwissen 2002), that have implemented the optimum genetic contribution theory (Woolliams & Thompson 1994) may be a way to achieving the operational objectives defined here.
- Minimize the co-ancestry among mated animals.
- The maximum number of living founders – or better, living animals born between the 1960s and 1970s or their direct offspring – should be represented in the next generation.
- Berga should import studs from remaining subpopulations, which should be as unrelated as possible to its jennies.

### Acknowledgements

This study was financed by the Departament d'Agricultura, Ramaderia i Pesca of the Generalitat de Catalunya (DARP). We would also like to thank the



AFRAC association for their helpful contribution and for providing the data used in the analysis.

## References

- Aranguren-Méndez J.A., Jordana J., Gómez M. (2001) Genetic diversity in Spanish donkey breeds using microsatellite DNA markers. *Genet. Sel. Evol.*, **33**, 433–442.
- Aranguren-Méndez J., Gómez M., Jordana J. (2002) Hierarchical analysis of genetic structure in Spanish donkey breeds using microsatellite markers. *Heredity*, **87**, 209–211.
- Aranguren-Méndez J.A., Beja-Pereira A., Avellanet R., Dzama K., Jordana J. (2004) Mitochondrial DNA variation and genetic relationships in Spanish donkey breeds (*Equus asinus*). *J. Anim. Breed. Genet.*, **121**, 319–330.
- Beja-Pereira A., England P.R., Ferrand N., Jordan S., Bakhiet A., Abdalla M.A., Mashkour M., Jordana J., Taberlet P., Luikart G. (2004) African origins of the domestic donkey. *Science*, **304**, 1781.
- Boichard D., Maignel L., Verrier E. (1997) The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.*, **29**, 5–23.
- Caballero A., Toro M.A. (2000) Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet. Res. Camb.*, **75**, 331–343.
- Caballero A., Toro M.A. (2002) Análisis de genética para el manejo de poblaciones subdivididas. *Conserv. Genet.*, **3**, 289–299.
- Dunner S., Checa M.L., Gutiérrez J.P., Martín J.P., Cañón J. (1998) Genetic analysis and management in small populations: the Asturcón pony as an example. *Genet. Sel. Evol.*, **30**, 397–405.
- Epstein H. (1984) Ass, mule and onager. In: I.L. Mason (ed.), *Evolution of Domesticated Animals*. Longman, London and New York, pp. 174–184.
- Falconer D.S., Mackay T.F.C. (1996) *Introduction to Quantitative Genetics*. Longman, Edinburgh.
- Folch P., Jordana J. (1998) Demographic characterization, inbreeding and maintenance of genetic diversity in the endangered Catalanian donkey breed. *Genet. Sel. Evol.*, **30**, 195–202.
- Goyache F., Gutiérrez J.P., Fernández I., Gómez E., Álvarez I., Diez 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**, 95–103.
- Gutiérrez J.P., Goyache F. (2005) A note on ENDOG: a computer program for analysing pedigree information. *J. Anim. Breed. Genet.*, **122**, 172–176.
- Gutiérrez J.P., Altarriba J., Díaz C., Quintanilla R., Cañón J., Piedrafita J. (2003) Pedigree analysis of eight Spanish beef cattle breeds. *Genet. Sel. Evol.*, **35**, 43–64.
- James J.W. (1972) Computation of genetic contributions from pedigrees. *Theor. Appl. Genet.*, **42**, 272–273.
- Jordana J., Folch P. (1998) La raza asnal catalana: programa de conservación y mejora de una población en peligro de extinción. *Arch. Zootec.*, **47**, 403–409.
- Jordana J., Folch P., Sanchez A. (1999) Genetic variation (protein markers and microsatellites) in endangered Catalanian donkeys. *Bioch. Syst. Ecol.*, **27**, 791–798.
- Jordana J., Folch P., Aranguren-Méndez J.A. (2001) Microsatellite analysis of genetic diversity in the Catalanian donkey breed. *J. Anim. Breed. Genet.*, **118**, 57–63.
- Kumar S., Tamura K., Jakobsen I.B., Nei M. (2001) MEGA 2: molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244–1245.
- Lacy R.C. (1989) Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zoo. Biol.*, **8**, 111–123.
- Meuwissen T.H.E. (2002) Gencont: an operational tool for controlling inbreeding in selection and conservation schemes. CD Proc. 7th World Congr. Genet. Appl. Livestock Prod. Montpellier, France, Communication No. 28–20.
- SAS/STAT (1999) SAS/STAT User's Guide. Release 8.2. SAS Institute Inc, Cary NC.
- Te Braake M.F.H., Groen A.F., Van Der Lught A.W. (1994) Trends in inbreeding in Dutch Black and White dairy cattle. *J. Anim. Breed. Genet.*, **111**, 356–366.
- Valera M., Molina A., Gutiérrez J.P., Gómez J., Goyache F. (2005) Pedigree analysis in Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. *Livest. Prod. Sci.* (in press).
- Woolliams J.A., Thompson R. (1994) A theory of genetic contributions. *Proc. 5th World Congr. Genet. Appl. Livestock Prod.*, **19**, 127–134.
- Wright S. (1931) Evolution in mendelian populations. *Genetics*, **16**, 97–159.
- Wright S. (1978) *Evolution and the Genetics of Populations: Vol. 4. Variability within and among Natural Populations*. University of Chicago Press, Chicago, IL, USA.