Genetic variation (protein markers and microsatellites) in endangered Catalonian donkeys

J. Jordana*, P. Folch, A. Sanchez

Unitat de Genètica i Millora Animal, Departament de Patologia i de Producció Animals, Facultat de Veterinària, Universitat Autònoma de Barcelona, 08193-Bellaterra, Barcelona, Spain

Received 17 September 1998; accepted 21 January 1999

Abstract

Genetic variation of the endangered Catalonian donkey breed (*Equus asinus*) has been analysed at 19 loci including seven protein loci and 12 microsatellite loci isolated from the domestic horse, in 98 individuals of both sexes. Only four protein markers and three microsatellites were polymorphic. Allele frequencies of the analysed loci showed close agreement with Hardy–Weinberg proportions, with the exception of the MPZ002 locus (*P* < 0.01). The within-population inbreeding estimate was not significantly different from zero (as measured by *F*<sub>IS</sub>-statistic). The cumulative-exclusion probability for all polymorphic loci was 82.9%, this value still being very low so that these markers could efficiently be utilised for verification of parentage. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Animal genetic resources; Donkey; Endangered breed; Genetic variation; Inbreeding; Parentage control

1. Introduction

The Catalonian donkey breed is a population in danger of extinction. It is found in several Pyrenean and pre-Pyrenean regions of the Catalonian area of Northeast Spain. At the last census there were only slightly more than 100 animals, a third of
which were males (Jordana and Folch, 1996). Following the guidelines proposed by
the FAO, this breed was characterised morphologically (Folch and Jordana, 1997),
haematologically (Folch et al., 1997), by clinical biochemical parameters (Jordana
et al., 1998) and demographically (Folch and Jordana, 1998). The purpose of this
paper is to genetically characterise this breed.

Markers used for such characterisation have been conventional biochemical poly-
morphisms (protein markers) and, more recently, DNA hypervariable sequences
known as microsatellites (SSR-simple sequence repeat or STR-short tandem repeat),
with these now being the markers of choice for the study of variation measurements
and genetic structure of populations (Goldstein and Pollock, 1997). There are only
a few such studies of *Equus asinus*, and these studies indicated a limited number of
polymorphic protein loci in the populations analysed (Niece and Kracht, 1967;
Mengozzi et al., 1982; Bowling and Nickel, 1985; Ketchum and Cothran, 1989; Bell,
1994; Ouragh et al., 1996); no study of microsatellite markers has been done in donkey
populations. Breen et al. (1994) in a study of only eight individuals, verified that a set
of 13 loci, isolated from the domestic horse (*Equus caballus*), could be amplified (by
PCR-polymerase chain reaction techniques) satisfactorily in donkeys.

The two main objectives of the Catalonian donkey breed Conservation Programme
are: (1) maintenance of maximum genetic diversity, and (2) minimization of any
consanguinity increase per generation. Thus, knowledge of the genetic population
structure is essential. Identification of polymorphic markers will allow us to describe
levels of genetic variability, estimate the degree of inbreeding, and identify the most
heterozygous individuals in the population in order to arrange the best matings (Gill
and Harland, 1992) for retaining the maximum ancestral genetic variability (Alderson,
1992; Boichard et al., 1997). Such markers will also provide reliable tools for parentage
verification, and will provide adequate information for phylogenetically relating this
breed to the other worldwide donkey breeds.

2. Materials and methods

2.1. Animals and genetic loci analysed

Blood samples from 98 individuals of both sexes, which represent approximately
95% of the total census of the breed, were examined for variation at seven protein loci
and 12 microsatellite loci. The 6-phosphogluconate dehydrogenase (PGD) and glu-
cose phosphate isomerase (GPI) red-blood-cell systems were analysed by horizontal
electrophoresis in agarose gel (Gahne and Juneja, 1985). The five plasma proteins
were: Albumin (ALB), detected by horizontal electrophoresis in starch gel (Bortolozzi,
1983); Transferrin (TF), Vitamin D-binding protein (GC), Alpha 1-β glycoprotein
(A1B) and Protease inhibitor (PI), typed by horizontal electrophoresis in polyacry-
lamide gels (Juneja and Gahne, 1987; Bell, 1994).

Twelve microsatellite loci isolated from the domestic horse, HMS1, HMS3, HMS5,
HMS6, HMS7, HTG6, HTG8, HTG14, HTG15, MPZ001, MPZ002 and VHL20
(Breen et al., 1994), were also analysed. Amplified PCR products were resolved using
10% PAGE and ethidium bromide staining. Two size markers were used: markers V and VIII from Boehringer (Boehringer Mannheim, Mannheim, Germany).

2.2. Statistical analyses

Alleles frequencies and mean heterozygosity values per polymorphic locus were obtained using BIOSYS-1 (Swofford and Selander, 1989). Tests of genotype frequencies for deviations from Hardy–Weinberg equilibrium were carried out using the exact tests of the GENEPOP computer programme (Raymond and Rousset, 1995). Tests were carried out separately for both types of genetic markers.

Using the methods of Weir and Cockerham (1984), as implemented in the FSTAT computer programme (Goudet, 1985), the \( f \)-statistic value for each locus was calculated. This statistic is analogous to Wright (1965, 1978) \( F_{IS} \)-statistic; i.e., it measures the deficit or the excess of heterozygotes. Significance was determined from permutation tests with the sequential Bonferroni procedure (Hochberg, 1988) applied over all loci (alleles were permuted within the population).

Polymorphic information content (PIC) for each microsatellite loci was calculated according to Botstein et al. (1980), and the probability of exclusion (PE) (Jamieson, 1994) was determined for all systems.

3. Results and discussion

Alleles frequencies, statistics of genetic variability, and PIC and PE values for biochemical and microsatellite loci are shown in Table 1.

Only four protein markers were polymorphic (TF, PGD, GC and PI). All allelic variants have been described previously, with a very similar distribution in their allelic frequencies (Bell, 1994). Of interest is the TF\( ^{D} \) variant, which was shown to be a rare allele (0.002–0.044) in several donkey populations from the USA, Australia, Italy and Morocco. In the Catalonian donkeys it had a higher value (0.272), possibly due to a bottleneck or founder event which this population might have experienced at some time in its history (Fig. 1). Only three of the microsatellite loci were polymorphic: HTG6, four alleles detected with sizes ranging from 82 bp to 90 bp; MPZ002, two alleles 80 to 89 bp; and VHL20, two alleles 86 to 104 bp (Fig. 2). The different variants found for each one of them (poly-, and monomorphic loci) are also within the ranges (bp size) described by Breen et al. (1994).

All protein markers showed close agreement with Hardy–Weinberg equilibrium (HWE). For the microsatellite loci, only MPZ002 showed significant disagreement with HWE, showing a very significant excess of heterozygotes (\( P < 0.01 \)).

When the population was treated as two subpopulations (males and females) no statistically significant differences in the allelic and genotypic distribution for any loci were observed (GENEPOP programme, data not shown). All loci showed agreement with HWE in both males and females, with the exception of the MPZ002 locus in females (\( P < 0.05 \)). The male subpopulation was in HWE for this locus (\( P > 0.05 \)), although barely so (\( P = 0.055 \)). These values confirm the excess of heterozygotes
Table 1
Allele frequencies, heterozygosity per locus and mean values ($H_o =$ observed; $H_e =$ expected), and exclusion probability (PE) per (a) analysed biochemical polymorphic locus, and (b) per analysed microsatellite polymorphic locus (PIC, polymorphic information content)

<table>
<thead>
<tr>
<th>Protein markers</th>
<th>Alleles</th>
<th>Frequency</th>
<th>$H_o$</th>
<th>$H_e^a$</th>
<th>PE</th>
<th>Microsatellite loci</th>
<th>Locus</th>
<th>Alleles</th>
<th>Frequency</th>
<th>$H_o$</th>
<th>$H_e^a$</th>
<th>PE</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>Ad</td>
<td>0.397</td>
<td></td>
<td></td>
<td></td>
<td>HTG6</td>
<td>A</td>
<td>0.092</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bd</td>
<td>0.217</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td>0.092</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cd</td>
<td>0.114</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>0.342</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dd</td>
<td>0.272</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td>0.474</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.783</td>
<td>0.712</td>
<td>0.464</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PGD</td>
<td>Cd</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
<td>MPZ002</td>
<td>A</td>
<td>0.520</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.979</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td>0.480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.042</td>
<td>0.041</td>
<td>0.020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>F</td>
<td>0.186</td>
<td></td>
<td></td>
<td></td>
<td>VHL20</td>
<td>A</td>
<td>0.429</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>0.814</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td>0.571</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.287</td>
<td>0.305</td>
<td>0.128</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>M</td>
<td>0.877</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0.126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.253</td>
<td>0.222</td>
<td>0.097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean heterozygosity</td>
<td></td>
<td>0.341</td>
<td></td>
<td>0.320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative exclusion probability</td>
<td></td>
<td>(0.157)</td>
<td></td>
<td>(0.142)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Unbiased estimate (Nei, 1978)
Standard errors in parentheses,
d = donkey specific variant.
observed for the global population, in males as well as females. The fact that MPZ002 genic frequencies are statistically equal for males and females (non-significant differences) made us reject the hypothesis of a possible Robertson Effect (Robertson, 1965) as an explanation for the significant excess of heterozygotes, perhaps indicating a locus-specific effect which suggests selection affecting this locus (Barker et al., 1997).
Fig. 2. Polymorphic microsatellite loci: VHL20, HTG6 and MPZ002, and their phenotypes. Amplified PCR products were resolved using 10% PAGE and ethidium bromide staining.

Table 2
$F_{IS}$-statistic analysis. Within-population inbreeding estimate ($f$) in the Catalonian donkey breed for both sets of markers (biochemical polymorphisms and microsatellite loci)

<table>
<thead>
<tr>
<th>Protein markers</th>
<th>$F_{IS} \equiv f$</th>
<th>Microsatellite loci</th>
<th>$F_{IS} \equiv f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus</td>
<td></td>
<td>Locus</td>
<td></td>
</tr>
<tr>
<td>TF</td>
<td>$-0.099$</td>
<td>HTG6</td>
<td>$0.002$</td>
</tr>
<tr>
<td>PGD</td>
<td>$-0.016$</td>
<td>MPZ002</td>
<td>$-0.304$</td>
</tr>
<tr>
<td>GC</td>
<td>$0.057$</td>
<td>VHL20</td>
<td>$0.005$</td>
</tr>
<tr>
<td>PI</td>
<td>$-0.139$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean estimate</td>
<td>$-0.079 (0.050)^a$</td>
<td>Mean estimate</td>
<td>$-0.086 (0.097)$</td>
</tr>
</tbody>
</table>

$^a$Mean estimate from jackknife over loci. Standard deviations in parentheses.

The within-population inbreeding estimates ($f \equiv F_{IS}$) are shown in Table 2. For each locus, no value was significantly different from zero. The mean estimates also were not significant. This $F_{IS}(f)$ average, obtained from jackknifing over loci, was equal to $-0.079 \pm 0.050$ for the protein markers. From a bootstrap analysis the true
value of the $F_{IS}$-statistic, with a 95% confidence interval, would range from $-0.120$ to $0.049$. For microsatellite loci, the mean estimate was equal to $-0.086 \pm 0.097$ and, with the same confidence interval, the range would be $-0.304$ to $0.005$.

These results indicate that there is not a significant degree of inbreeding in this population. To obtain more reliable inbreeding values, additional loci need to be analysed. However, these results do allow us to postulate that inbreeding, if it occurs, is negligible. It is well known that inbreeding affects all or most loci in a similar way with respect to heterozygotes. In the present work no deficit for any of the seven polymorphic markers analysed has been detected. So, although the population of Catalonian donkeys is classified as a breed in danger of extinction with the associated problems of increased consanguinity, these results indicate that the breeding policy in the last few decades has successfully avoided mating among closely related individuals to the maximum extent possible.

Table 1 shows the theoretical exclusion probabilities (PE) for each polymorphic locus, ranging from between 0.020 for the PGD system to 0.464 for the TF locus, with the cumulative PE for all loci being 0.829. This 82.9% of PE is too low to be useful as an effective tool in parentage verification. This control is very important for this endangered breed. The optimal mating being pursued by the Programme of Conservation’s technicians is between a stallion and a jenny, which minimizes the inbreeding coefficient and maximizes the genetic ancestral variability retention of the hypothetical offspring of the pair (Folch and Jordana, 1998).

Acknowledgements

The authors thank the Departament d’Agricultura, Ramaderia i Pesca of the Generalitat de Catalunya, which financed this study; and the AFRAC association for their helpful co-operation and assistance during sample collection. We also thank Chuck Simmons for the English revision.

References


Juneja, R.K., Gahne, B., 1987. Simultaneous phenotyping of pig plasma α-protease inhibitors (PI1, PO1A, PO1B, P12) and four other proteins (PO2, TF, CP, HPX) by a simple method of 2D horizontal electrophoresis. Anim. Genet. 18, 197–211.


