Inbreeding and Genetic Structure in the Endangered Sorraia Horse Breed: Implications for its Conservation and Management

CRISTINA LUIS, E. GUS COTHNAN, AND MARIA DO MAR OOM

From the Departamento de Biologia Animal, Faculdade de Ciências, Centro de Biologia Ambiental, Universidade de Lisboa, Edificio C2-Piso 3, Campo Grande, 1749-016 Lisboa, Portugal (Luís and Oom); and the Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843 (COTHNAN).

Address correspondence to C. Luís at the address above, or e-mail: cmluis@fc.ul.pt.

Abstract

The Sorraia horse is a closed breed with reduced effective population size and considered in critical maintained risk status. The breed exists in 2 main breeding populations, one in Portugal and one in Germany, with a smaller population size. A set of 22 microsatellite loci was used to examine genetic diversity and structure of the Sorraia horse breed and to compare individual inbreeding coefficient $F$, estimated from pedigree data, with individual heterozygosity and mean $d^2$. The Sorraia horse shows lower levels of microsatellite diversity when compared with other horse breeds. Due to management strategies, there are clear differences in the genetic structure of the 2 main Sorraia horse populations. Individual heterozygosity was shown to be a good estimator, used together with or as an alternative to inbreeding coefficient, in predicting fitness and evaluating the inbreeding level of the Sorraia horse. The information gathered in this study, combined with information available from previous studies, offers an important and wide information base for the future development of an effective breeding management of the Sorraia horse in order to preserve this endangered breed.

In 1975, 2 other farms started breeding the Sorraia horse, and in 1976, a subpopulation was created in Germany with 3 stallions and 3 mares imported from Portugal with no further immigration until the recent past (Oom 1992; Oom and Luís 2001). Although there are now other breeders in Germany, these animals are the founders of the majority of the population bred outside Portugal.

The Sorraia horse is a closed breed, with complete pedigree for all individuals tracing back to the founders. In 2004, the first volume of the Sorraia Studbook was published (Oom et al. 2004) containing 564 animals from the foundation of the breed until 31 December 2002. Less than 200 individuals are alive, and only 10 of the 12 founders are still represented in the living population (Oom et al. 2004; Kjollerstrom 2005). The Sorraia breed has a reduced effective population size and, according to Food and Agriculture Organization (FAO) criteria, is considered in critical maintained risk status (FAO 1998; DAD-IS Database http://dad.fao.org). The low population size has led to high levels of mean kinship (mk) coefficient ($0.282 \leq mk \leq 0.423$) and average inbreeding...
(F = 0.363), as estimated from genealogical data analysis (Kjøllerström 2005; Oom and Serrano 2006).

Compelling reasons for conserving genetic resources of domestic livestock, including horses, can be classified into a few categories that include genetic insurance for the species and for agriculture, historical and cultural factors, and concerns for the extinction of variants that might have scientific value (Spönenberg 2000). The main objective of conservation genetics is to preserve variability within populations under the hypothesis of correlation between genetic variation and population viability (Hall and Bradley 1995). Studying homozygosity as an index for inbreeding is a valuable management tool and may be critical for the Sorraia horse.

Assessment of genetic variability in populations has been greatly facilitated by the advent of high-resolution molecular genetic techniques for population genetic studies, thus complementing the evaluation of genetic diversity by pedigree data analysis. Because of their high variability, microsatellites are often used to determine the amount and pattern of genetic variation in populations (Goldstein and Schlottner 1999). Using microsatellite loci information, some studies have estimated individual heterozygosity and mean d² to measure the extent of inbreeding in populations (e.g., Colman et al. 1998; Coulson et al. 1998; Ellegren 1999; Hedrick et al. 2002; Luís et al. 2003; Markert et al. 2004).

With only 12 founders and no further genetic input into the population since its foundation, the Sorraia horse shows limited genetic variability across several genetic markers, most likely due to founder effect and genetic drift as detected in several previous works (Oom and Cothran 1994; Luís, Bastos-Silveira, et al. 2002; Luís, Cothran, et al. 2002; Luís et al. 2005). There is also a significant indication that inbreeding levels in the Sorraia horse are negatively correlated with both male and female fertility rates and with juvenile survival (Kjøllerström J, do Amaral JP, do Mar Oom M, in preparation). Those problems may be of particular concern as such a small effective population size makes the population even more sensitive to demographic stochasticity.

The aim of this study is to use a set of microsatellite data to examine genetic diversity and structure of the Sorraia horse breed and to compare individual inbreeding coefficient $F_1$, estimated from pedigree data, with individual heterozygosity and mean $d^2$. With this study, we intend to gather genetic information that may complement the pedigree information and be useful for managing the genetic diversity of the endangered Sorraia breed.

**Materials and Methods**

**Sampling and DNA Extraction**

A total of 131 Sorraia horses, used to belong to the 2 main populations of this breed, were used as source material for DNA extraction: 99 individuals from Portugal (including individuals from 3 main breeding farms) and 32 individuals from Germany (including the 5 founders of this population). DNA was extracted from fresh whole-blood samples using the high-salt extraction procedure (Montgomery and Sise 1990) and from frozen samples with the QIAamp DNA Blood Mini Kit (Qiagen Inc., Valencia, CA).

**Microsatellite Analysis**

Twenty-two dinucleotide horse microsatellites, distributed over 16 different autosomes (Table 1), were analyzed: \( AHT4, AHT5, AHT7, AHT8, AHT9 \) (Binns et al. 1995), \( ASB2 \) (Breen et al. 1997), \( HMS1, HMS2, HMS3, HMS6 \) (Guérin et al. 1994), \( HTG4, HTG6, HTG8 \) (Ellegren et al. 1992), \( HTG10, HTG14 \) (Marklund et al. 1994), \( LEX20, LEX23 \) (Google et al. 1996), \( LEX36, LEX41 \) (Google et al. 1997), \( VHL20 \) (Van Haeringen et al. 1994). Polymerase chain reactions (PCRs) and amplification were performed as described by Luís, Cothran, et al. (2002).

PCR products were separated by electrophoresis in a 6% Long Ranger gel solution (Cambrex Bio Science, Rockland, ME) using an automated fluorescence Li-Cor 4200S sequencer (Li-Cor, Lincoln, NE). Alleles were scored according to PCR product size with RFLPscan 3.1 software (Scanalytics CPS Inc., Rockville, MD).

**Data Analysis**

The number of individuals typed, mean number of alleles (MNA), the observed heterozygosity (Ho), and the unbiased expected heterozygosity (He) were estimated with the software package GENETIX v. 4.04 (Belkhir et al. 2002). Polymorphic information content (PIC), and the probability of exclusion (PE) were assessed using the computer program CERVUS (Marshall et al. 1998).

The software GENEPOP v. 3.3 (Raymond and Rousset 1995) was used to test for deviations from Hardy–Weinberg equilibrium (HWE) across all loci and subpopulations applying the exact test with default settings of the Markov chain Monte Carlo methodology (Guo and Thompson 1992). P values were adjusted by applying the Hochberg’s step-up Bonferroni method (Hochberg 1988).

$F_1$ statistics of Wright (1951) were used to analyze genetic structuring of the Sorraia breed. The fixation coefficients ($F_{ST}$ and $F_{IS}$) were estimated according to Weir and Cockerham (1984) using the FSTAT 2.9.3.2 computer program (Goudet 2002).

The existence of genetic bottlenecks was tested using the program BOTTLENECK (Cornuet and Luikart 1996). In recently bottlenecked populations, the majority of loci will exhibit an excess of heterozygotes, thus exceeding the heterozygosity expected in a population at mutation drift equilibrium. The excess of such heterozygosity was estimated under the assumption of a two-phase microsatellite mutation (TPM) model (Luikart et al. 1998) and under application of the Wilcoxon signed-rank test. BOTTLENECK was further employed to test for an L-shaped distribution of allele frequencies of each marker locus because such a distribution is not typical for populations that have undergone short but severe bottleneck (Luikart et al. 1998).
Individual heterozygosity was calculated as the number of loci at which an individual was heterozygous, divided by the total number of loci at which an individual was scored. Mean $d^2$ was calculated as the squared distance in repeat units between the 2 alleles an individual had at a microsatellite locus, averaged over all loci at which an individual was scored, according to equation:

$$\text{mean } d^2 = \frac{1}{n} \sum_{i=1}^{n} (i_a - i_b)^2,$$

where $i_a$ and $i_b$ are the lengths (in base pair) of alleles $a$ and $b$ at locus $i$ and $n$ is the total number of loci at which an individual was scored (Coulson et al. 1998).

Because not all individuals were successfully scored at all loci, individuals that had been scored at fewer than 11 loci were omitted from individual heterozygosity and mean $d^2$ analyses.

The inbreeding coefficients values used are included in the Sorraia Studbook (Oom et al. 2004) and were estimated according to Oom (1992) with all genealogies tracing back to the founders.

Descriptive statistics and the Pearson’s correlation between individual heterozygosity, mean $d^2$, and inbreeding coefficient variables were calculated using SAS® software package (RELEASE 8.2; SAS Institute Inc., Cary, NC, 2003). Pearson’s correlation was used because the variables have a normal distribution.

## Results

### Genetic Diversity

The total number of alleles found for the 22 microsatellites analyzed was 73, and the MNA per locus was 3.3, with a range of 2–5 alleles. Ho ranged between 0.088 (HMS2) and 0.705 (LEX36). The highest values for He, PIC, and PE were found for microsatellite HTG4, whereas the lowest ones were found for microsatellite HMS2 (Table 1).

After Hochberg’s step-up Bonferroni correction for multiple comparisons (Hochberg 1988), a significant deviation from HWE was only observed for microsatellite HTG10. A positive nonsignificant $F_{IS}$ estimation ($P > 0.05$) was obtained for the Portuguese population, whereas a highly significant negative estimation ($P < 0.001$) was obtained for Germany.

For the Portuguese population, under the assumption of the TPM, the Wilcoxon test gave a significant result ($P = 0.0165$) indicative of a bottleneck. However, the presence of an L-shaped distribution of the allele frequencies does not indicate a recent bottleneck in this population. For the German population, the Wilcoxon test revealed the existence...
of a highly significant bottleneck \((P \leq 0.001)\) that was confirmed by the absence of an L-shaped distribution of the allele frequencies.

Population Structure

The pairwise \(F_{ST}\) coefficient, representing the differentiation of the 2 main Sorraia populations, was highly significant with a value of 0.09 \((P \leq 0.001)\). Thus, 9\% of the microsatellite variability is explained by the subdivision of populations, whereas the remaining variability is explained by the variation within populations.

Inbreeding, Individual Heterozygosity, and Mean \(d^2\)

Frequency distributions of the mean \(d^2\), individual heterozygosity, and inbreeding coefficient are shown in Figure 1. Mean individual heterozygosity and mean \(d^2\) were 46.4 ± 11.3\% and 36.9 ± 16.7\% (in bp\(^2\)), respectively, whereas mean inbreeding coefficient for the analyzed samples was 32.5 ± 7.2\%. There was a highly positive correlation \((r = 0.395, P \leq 0.0001)\) between individual heterozygosity and mean \(d^2\). A nonsignificant negative correlation was observed between inbreeding coefficient and mean \(d^2\) \((r = -0.116, P = 0.198)\), whereas a very significant negative correlation was observed between inbreeding coefficient and individual heterozygosity \((r = -0.268, P < 0.005)\).

Discussion

The Sorraia horse shows lower levels of microsatellite diversity when compared with the majority of other horse breeds (e.g., He: 0.677–0.770, Ho: 0.694–0.765, MNA: 5.2–7.8, Cañón et al. 2000; He: 0.646–0.732, Ho: 0.628–0.671, MNA: 4.7–7.5, Cunningham 2001; He: 0.751, Ho: 0.732, MNA: 4.5, Morais et al. 2005; He: 0.442–0.770, Ho: 0.452–0.785, MNA: 2.0–4.7, Juras et al. 2003; He: 0.675, Ho: 0.663, MNA: 7.1, Achmann et al. 2004; He: 0.650–0.765, Ho: 0.633–0.777, MNA: 4.33–8.08, Solis et al. 2005). The German and Portuguese subpopulations differ from each other in variability. As expected, due to founder effect, the German population has a reduced number of alleles for most loci analyzed compared with the Portuguese one; however, no loss of founders’ alleles at tested loci has occurred since the foundation of this population due to genetic drift. Both Ho and He were higher in the German population, despite statistical evidence for the genetic bottleneck that took place in 1976. The lower heterozygosity found in Portugal is most probably caused by the breeding management used in most Portuguese breeding farms with breeding done in pasture by free live-cover, and the use of only 1 stallion per year per herd, with this same stallion often used during several breeding years. This strategy is completely different from the one used in Germany, where breeding is mostly done by hand-cover with different males used almost each year and more than one per herd, a management strategy that appears to be more suited for genetic conservation because in this population the genetic variability parameters are higher.

As expected, the significant \(F_{ST}\) coefficient estimated indicates a relatively low gene flow between the Portuguese and German subpopulations sampled due to the reproductive isolation that existed between them until very recently when some Sorraia from Portugal were taken to Germany by new German breeders. Thus, the management strategies implemented in the 2 populations seem to be the main reason for the existence of different genetic structures in both populations.

Founder effect, genetic drift, and artificial selection/breeding management are among the factors that likely contributed to the differentiation within the Sorraia horse breed, owing to the low size of the analyzed populations. Migration can be considered null because there has been no migration between populations since fragmentation, at least until very recently and after sampling collection for this study was completed. Also, in most domestic species, the drift process can be accelerated by the unequal representation of both sexes, which is the case of the Sorraia where there are far fewer males in the breeding population.
A major conclusion from this study related to genetic conservation in the Sorraia breed is that the Portuguese breeders should try to focus the breeding management strategy in increasing heterozygosity. Some efforts have been made, over the last 3 breeding years, by the Sorraia Breeder’s Association to advise Portuguese breeders of the best suitable stallions to cover the mares from different herds, based on pedigree information and subsequent genetic analysis performed with GENES software v. 11.8 (Lacy 1998). It will be interesting to analyze the results of this practice in the future in terms of molecular genetic variation within the Portuguese population and the Sorraia total population.

Following a strategy that is fully recommended in conservation plans, in recent years, and after the sampling for this study was performed, the Portuguese population has been further fragmented into different and smaller subpopulations. The intention is to decrease the loss of the genetic variability per generation of the overall population by promoting a more effective exchange of stallions between subpopulations (which will undergo different local effects of genetic drift) and to reduce kinship between animals of different subpopulations.

Pedigrees are essential to provide information for breeding management plans, on the assumption that data registered is correct. A panel of 15 microsatellites showing a PE of 99.8% is now used for parentage testing in the Sorraia horse. The set of microsatellites used in this study would provide superior power of exclusion and thus would provide a superior panel for paternity testing of the Sorraia horse.

For the Sorraia horse, a population that is subject to close inbreeding, individual heterozygosity can more effectively be used to evaluate inbreeding level of the population and, inbreeding depression. Therefore, within the Sorraia breed, this molecular estimator could be used, together with, or alternatively to inbreeding coefficient, in predicting fitness. The results of this work support the findings by Tsitrone et al. (2004), stating that heterozygosity is a more robust measure of inbreeding than is mean inbreeding coefficient, in predicting fitness. Therefore, within the Sorraia breed, this molecular estimator could be used, together with, or alternatively to inbreeding coefficient, in predicting fitness.

The results from this work support the findings by Tsitrone et al. (2001), Slate and Pemberton (2002), and Markert et al. (2004) stating that heterozygosity is a more robust measure of inbreeding than is mean I and that multilocus heterozygosity data can substitute or provide a useful complement to data on inbreeding depression obtained through pedigree analysis.

The results of the present study are in accordance with the known history of the Sorraia horse breed that is a highly inbred, small, and closed population. As stated by Fernández et al. (2004), information obtained from the analysis of neutral molecular markers can be used to make decisions about individuals’ contributions to the next generation in order to maximize gene or allelic diversity. Therefore, the information gathered in this study, combined with information available from pedigree and reproductive fitness analysis (Oom 1992; Matos 1996; Oom and Luí 2001; Kjøllerstrøm 2005), from morphological measures (Oom 1992), from mitochondrial DNA markers (Luís, Cothran, et al. 2002; Luís, Bastos-Silveira, Cothran, and Oom 2006; Luís, Bastos-Silveira, Cothran, Costa-Ferreira, et al. 2006), and immunogenetics (Luís et al. 2005), offers an important and wide information base for the future development of effective breeding management of the Sorraia horse in order to preserve this endangered breed.

Acknowledgments

We would like to thank Joana Morais, Ana Rina Grosso, and Joana Cavaco Silva for help provided in the laboratory. C.L. was supported by a PhD grant (SFRH/BD/3318/2000) from the Portuguese Foundation for Science and Technology (FCT/MCT).

References


d’Andrade R. 1926. Apontamentos para um estudo sobre a origem e domes- ticac do cavalo na Penı ´nsula Ibe´rica. Lisboa (Portugal): Aproximac oes, Centro Tipográfico Colonial.


d’Andrade R. 1954. Alredor del Caballo Espan˜ol. Lisboa (Portugal): Colec- 

Downloaded from https://academic.oup.com/jhered/article-abstract/98/3/232/2187893 by guest on 21 March 2019


Goudet J. 2002. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2) [Internet]. Available from: http://www2.unil.ch/popgen/softwares/fstat.htm.


Received August 7, 2006
Accepted February 12, 2007

Corresponding Editor: Ernest Bailey