doi:10.1111/j.1365-2052.2005.01376.x

Genetic diversity among horse populations with a special focus on the Franches-Montagnes breed

M. L. Glowatzki-Mullis*, J. Muntwyler*, W. Pfister*, E. Marti†, S. Rieder‡, P. A. Poncet§ and C. Gaillard*

*Institute of Animal Genetics, Nutrition and Housing, Vetsuisse Faculty, University of Berne, CH-3001 Berne, Switzerland. †Division of Clinical Immunology, Vetsuisse Faculty, University of Berne, CH-3001 Berne, Switzerland. ‡Swiss College of Agriculture, CH-3052 Zollikofen, Switzerland. §Haras National Suisse, CH-1580 Avenches, Switzerland

Summary

Genetic characterization helps to assure breed integrity and to assign individuals to defined populations. The objective of this study was to characterize genetic diversity in six horse breeds and to analyse the population structure of the Franches-Montagnes breed, especially with regard to the degree of introgression with Warmblood. A total of 402 alleles from 50 microsatellite loci were used. The average number of alleles per locus was significantly lower in Thoroughbreds and Arabians. Average heterozygosities between breeds ranged from 0.61 to 0.72. The overall average of the coefficient of gene differentiation because of breed differences was 0.100, with a range of 0.036–0.263. No significant correlation was found between this parameter and the number of alleles per locus. An increase in the number of homozygous loci with increasing inbreeding could not be shown for the Franches-Montagnes horses. The proportion of shared alleles, combined with the neighbour-joining method, defined clusters for Icelandic Horse, Comtois, Arabians and Franches-Montagnes. A more disparate clustering could be seen for European Warmbloods and Thoroughbreds, presumably from frequent grading-up of Warmbloods with Thoroughbreds. Grading-up effects were also observed when Bayesian and Monte Carlo resampling approaches were used for individual assignment to a given population. Individual breed assignments to defined reference populations will be very difficult when introgression has occurred. The Bayesian approach within the Franches-Montagnes breed differentiated individuals with varied proportions of Warmblood.

Keywords assignment, clustering, Franches-Montagnes, genetic diversity, horse, microsatellites.

Introduction

The Franches-Montagnes, the only native Swiss horse breed, was developed in the 19th century from basically local horses, some Comtois mares, and some Anglo-Norman sires. To satisfy transportation, agriculture and military needs, a few European Warmblood and Coldblood breeds were introgressed into the Franches-Montagnes during the first half of the 20th century. The mechanical revolution shifted traditional usage to other purposes, such as sports, leisure activities and hippotherapy. Some restricted introgression of European Warmblood and Arabian was practiced in the second half of the 20th century in response to the demand for a lighter horse type. This introgression was applied only with stallions that met the standards of the Franches-Montagnes. Since 1997, introgression with Warmblood and with other breeds has been prohibited. Mechanization led to a dramatic decrease in the horse population, but their numbers have recovered slightly in the last decade. The Franches-Montagnes breed includes about 3500 foaling mares and 150 stallions (a list of domestic animals is available at http://www.fao.org/dad-is/).

Within the framework of breed conservation, genetic characterization is important in guarding breed integrity and is a prerequisite for managing genetic resources

Address for correspondence

Marie-Louise Glowatzki-Mullis, Clinical Research, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Berne, Bremgartenstrasse 109a, CH-3001 Berne, Switzerland.
E-mail: marie-louise.glowatzki@itz.unibe.ch

Accepted for publication 24 August 2005

(Bjørnstad & Røed 2002). Microsatellites are widely adopted to quantify genetic variation within and among breeds and to assign individuals to reference populations (Bruford *et al.* 2003; Rosenberg *et al.* 2003). In this study, 50 microsatellite markers were used to investigate genetic divergence between six horse breeds: four represented populations that have been isolated for a long time (Thoroughbred, Arab, Icelandic Horse and Comtois), and two represented different degrees of outbreeding (Franches-Montagnes and European Warmblood). Special emphasis was placed on measuring the degree of introgression of European Warmblood into the Franches-Montagnes breed. Genetic diversity was measured by two clustering methods. The reliability of assigning or excluding an individual from its population was also investigated.

Materials and methods

Sampling and sample preparation

Blood or hair samples of 319 horses from the following populations were collected: Comtois (COM) 33, Icelandic Horse (ICL) 33, European Warmblood (WB) 50, Thoroughbred (TB) 31, Arabian (AR) 36 and Franches-Montagnes (FM) 136. Inbreeding coefficients were available on 107 FM animals based on pedigrees with at least 10 complete generations. DNA was extracted using a salting-out procedure.

Microsatellite analysis

A total of 52 microsatellite markers with known genomic assignments were included in this study. All chromosomes except the Y chromosome were represented. Four multiplex polymerase chain reactions (PCRs) using fluorescently labelled primers were developed (17plex: *ASB2*, *EB2E8*, *HMS1*, *HMS5*, *HTG3*, *HTG6*, *HTG7*, *TKY294*, *TKY297*, *TKY301*, *TKY321*, *TKY333*, *TKY337*, *TKY374*, *TKY394*, *UM11*, *UM32*; 14plex: *AHT4*, *AHT5*, *ASB17*, *ASB23*, *COR18*, *COR58*, *HMS3*, *HMS6*, *HMS7*, *HTG4*, *LEX33*, *TKY312*, *TKY343*, *VHL20*; 13plex: *AHT31*, *ASB43*, *COR22*, *COR69*, *COR7*, *HMS2*, *HTG10*, *I-18*, *LEX54*, *LEX73*, *LEX78*, *SGCV28*, *UCDEQ(CA)425* and 8plex: *COR55*, *LEX3*, *TKY279*, *TKY287*, *TKY325*, *TKY341*, *TKY344*, *UM5*). Detailed marker information is available in Table S1. PCR amplification was performed in a total volume of 15 μl with 10–50 ng template DNA, 2.1 μl GeneAmp 10x PCR Buffer I, 4.5 μl dNTPs 1.25 mM, primer mix and 0.4 μl AmpliTaq Gold DNA Polymerase. The cycling conditions included an initial activation step at 94 °C for 12 min, 30 cycles of 94 °C for 1 min (ramp 1 min), annealing at 58 °C for 1 min (ramp 1 min) and extension at 72 °C for 1 min (ramp 1 min) and a final extension at 72 °C for 45 min. Amplification was carried out using a PE GeneAmp PCR 9600 or 9700 system (Applied Biosystems, Foster City, CA, USA). PCR products were diluted with 80 μl distilled water if the amplified DNA was initially extracted from blood or with 20 μl distilled water if the amplified DNA was initially extracted from hair.

Detection of amplified products

Each diluted PCR reaction (1.2 μl) was mixed with 0.4 μl GeneScan 500 LIZ Size Standard and 10.6 μl Hi-Di formamide. The denatured samples were run (POP-4; run temperature 45°C) on a ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). Data collection, extraction and analysis were performed with 3100 Data Collection Software Version 1.0.1 and GeneScan® Analysis Software Version 3.7 (Applied Biosystems). Alleles were assigned using Genotyper® Software Version 3.7 from Applied Biosystems. Samples from the 2001 International Society for Animal Genetics comparison test (http://www.isag.org.uk/journal/comparisonguide.asp) were used as reference samples to standardize allele sizes.

Data analysis

The GENEPOP 3.4 program (Raymond & Rousset 1995) was used to calculate the allele frequencies and to test for Hardy–Weinberg equilibrium. The unbiased coefficient of gene differentiation, G_{ST} (Nei 1973; Nei & Chesser 1983), was computed using GENETIX (Belkhir *et al.* 1996–2002). The natural logarithm of the proportion of shared alleles (Bowcock *et al.* 1994) was chosen as the distance between individuals and calculated by MICROSAT 1.5b (Minch *et al.* 1998). The neighbour-joining method (Saitou & Nei 1987) was used to build a tree from the genetic distance matrix using the PHYLIP program package (Felsenstein 1993–2002). The results were entered into TreeExplorer (http://evolgen.biol.metro-u.ac.jp/TE/TE_man.html) in order to find a suitable graphic display.

Based on genotypes at marker loci, individuals were clustered into a given number of populations and assigned probabilistically to clusters inferred with the Bayesian approach of STRUCTURE (Pritchard *et al.* 2000). This approach is based on an admixture model in which an individual may have mixed ancestry. The model estimates the posterior mean probability that an individual multilocus genotype originates from populations among the inferred clusters. To choose the appropriate number of inferred clusters to model the data, three to nine inferred clusters were performed with five independent runs each. All analyses used a burn-in period of 10 000 and 100 000 iterations for data collection. Results were entered into DISTRUCT (Rosenberg 2004) to develop a graphic display.

The probability for each individual assigned to one or more of the defined populations was estimated using a Monte Carlo resampling method (Paetkau *et al.* 2004) implemented in GeneClass2, applying 10 000 resampling and a type I

error of 0.01 (Piry *et al.* 2004). Significant differences between the various means for heterozygosity and number of alleles per locus were tested according to Holm's test statistic (Proc MULTTEST, SAS® software package release 9.1, 2004; SAS Institute Inc., Cary, NC, USA).

Results

Microsatellite loci

All 52 microsatellites except *EB2E8* (occasional missing alleles) amplified well in the four multiplex PCR systems, so *EB2E8* was excluded from the data analysis. Because *LEX3* is located on the X chromosome and gender was only known for two breeds, this microsatellite was not included in the data analyses.

The total number of alleles was 402. The average number of alleles per locus was 8.0 (Table S2), ranging between 3 (*HMS5*) and 15 (*ASB17*). The average number of alleles per locus for each breed varied between 4.7 (TB) and 6.6 (WB) (Table 1). The average number in FM was 6.4; for AR and TB this number was significantly lower than for FM, ICL and WB ($P < 0.05$). The average expected and observed heterozygosities were very similar within breeds and ranged from 0.61 (AR) to 0.72 (WB) across breeds. The expected heterozygosities between AR and WB ($P = 0.001$), as well as the observed heterozygosities between AR and WB and between AR and FM ($P = 0.013$ and $P = 0.023$ respectively) differed significantly.

All six populations were in Hardy–Weinberg equilibrium. Breed-specific alleles with a frequency of over 0.20 were observed in FM and ICL. The unbiased coefficients of gene differentiation (G_{ST}) due to breed differences ranged from 0.036 (*HMS1*) to 0.263 (*I-18*), with an overall average of 0.100 (Table S2). The correlation between the average number of alleles and G_{ST} equalled −0.023.

No significant increase of the number of homozygous microsatellite loci was observed with increasing inbreeding coefficients in FM (regression coefficient = −0.054, $P = 0.750$) (Fig. S1).

Breed differentiation

The neighbour-joining dendrogram of the allele-sharing distance (Fig. 1) formed clearly defined clusters for FM, COM, ICL and AR; only one COM was misplaced. Most individuals of TB and WB were closely assembled, but there was no definite differentiation between these two breeds. Moreover, three WB animals clustered with AR.

Clustering using a Bayesian approach was performed on the entire data set with increasing numbers of inferred clusters. The most appropriate number of clusters for modelling the data were eight. When the number of clusters was fixed at six, the six breeds each formed a separate cluster. A new cluster of horses appeared within WB when the number of clusters was fixed at seven. Another new cluster could be seen within FM when eight clusters were inferred (Fig. 2, upper panel). AR, COM, ICL and TB were each grouped in their own cluster with an estimated membership of ≥0.95 (clusters VII, VIII, V and II respectively; Table 2). FM horses were mainly assigned to cluster IV (estimated membership 0.86) and to a small extent to cluster III (estimated membership 0.09). Cluster III characterizes WB within the FM, but these WB were almost not present in the analysed WB (estimated membership 0.009). The estimated membership of FM in clusters representing breeds that were involved in the build-up of FM (AR, COM, WB) was insignificant (≤0.01). The WB were split in the three clusters VI, I and II with an estimated membership of 0.50, 0.32 and 0.13 respectively. An additional clustering with the Bayesian approach was performed with FM in order to obtain more information about the population structure. The results with four clusters are shown in Fig. 2, lower panel. The average proportion of Warmblood in the four clusters was 0.03, 0.06, 0.13 and 0.29. Differences between these proportions were significant ($P < 0.0001$), except between clusters A and B ($P = 0.164$), but in B one founder stallion occurred more frequently as ancestor than in A. Two other founder stallions that were used simultaneously 15 years ago predominantly appeared in clusters C and D.

Table 1 Heterozygosity, average number of alleles and breed-specific alleles in six horse breeds.

	Heterozygosity		Average number of alleles per locus (SE)	Breed-specific alleles with frequency ≥20%		
Breed[1]	Expected (SE)	Observed (SE)		Locus	Allele[2]	Frequency (%)
AR	0.63 (0.02)	0.61 (0.02)	4.8 (0.20)			
COM	0.68 (0.02)	0.67 (0.02)	5.6 (0.21)			
FM	0.69 (0.01)	0.70 (0.01)	6.4 (0.26)	*TKY287*	u	25
ICL	0.69 (0.02)	0.69 (0.02)	5.9 (0.23)	*COR18*	s	20
WB	0.72 (0.01)	0.71 (0.01)	6.6 (0.26)			
TB	0.67 (0.01)	0.65 (0.02)	4.7 (0.17)			

[1]AR, Arabian; COM, Comtois; FM, Franches-Montagnes; ICL, Icelandic Horse; WB, European Warmblood; TB, Thoroughbred.
[2]Allele u, *TKY287* ~ 243 bp, allele s, COR18 ~ 275 bp (more information in Table S1).

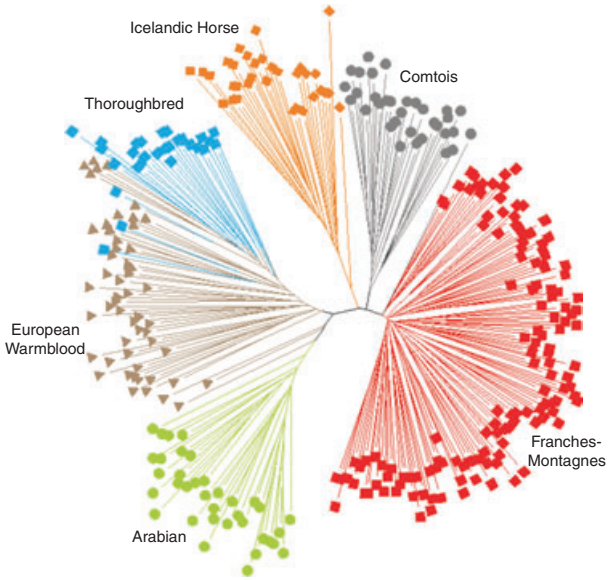

Figure 1 Neighbour-joining tree summarizing the allele-sharing distances among 319 horses based on 50 microsatellites.

Assignment of individuals to pre-defined populations

In 94–100% of the cases, individuals could be assigned to a pre-defined population (Table 3; Table S3). However, a large number of AR and TB animals were not only correctly assigned to the pre-defined breed but also to WB (78% and 100% respectively). Three per cent of the AR horses were falsely assigned as WB. A few COM, FM and ICL individuals could not be assigned to any of the six breeds (6%, 1% and 3% respectively).

Discussion

The number of alleles per locus is frequently used as a measure of polymorphism, and loci with a low number of alleles are often discarded from analyses. However, according to our findings there was no relationship between the number of alleles and the coefficient of gene differentiation, G_{ST}. For example, *HMS5*, with only three alleles, had a G_{ST} value of 0.068 across six breeds, which was similar to the G_{ST} value of 0.065 for *UM11* with 14 alleles. The correlation between the number of alleles per locus and G_{ST} was not significant (−0.02), and it was very similar (−0.09) to the correlation that we obtained with the data of Bjørnstad *et al.* (2000).

Of 50 microsatellites, only two breed-specific alleles were found with a high frequency (>20%). This low number is not surprising, because outbreeding has been practiced in horse breeding, especially in WB with TB. Furthermore, the time since the formation of the breeds as closed populations is still too short for genetic drift to have had an effect.

These 50 microsatellites explain about 10% of the genetic difference between the six breeds, which is similar to the G_{ST} values of 8% in the study of Cañon *et al.* (2000) and 11.6% for heavy horse breeds in the study of Aberle *et al.* (2004); Bjørnstad *et al.* (2000) obtained a somewhat higher value

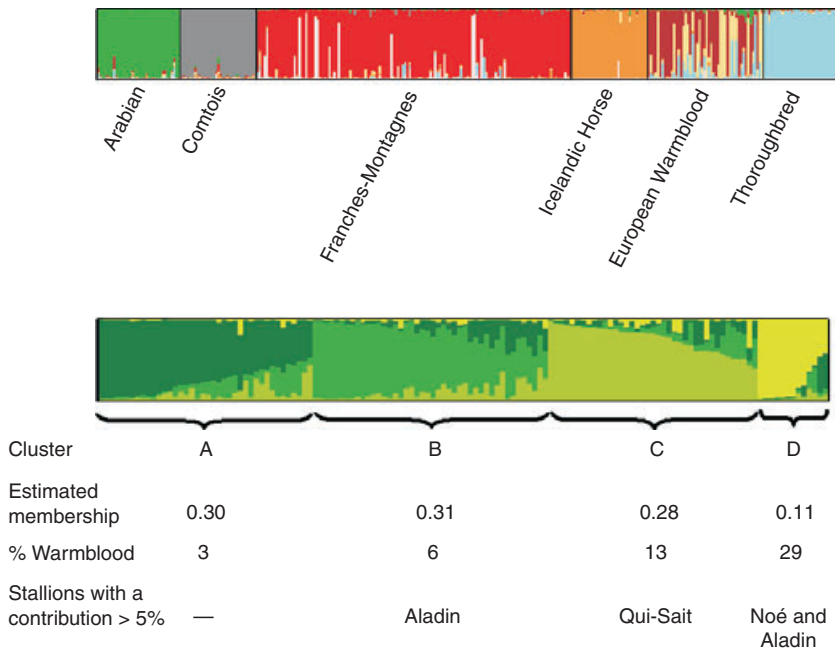

Figure 2 Population structure obtained by STRUCTURE analyses. (upper panel) Proportion of membership of six horse breeds in eight inferred clusters. Each of the 319 horses is represented by a thin vertical line that is divided into coloured segments representing the individual's membership in the cluster of the corresponding colour. (lower panel) Proportion of membership of 136 Franches-Montagnes horses in four inferred clusters representing different degrees of introgression with Warmblood.

Table 2 Assignment of 319 horses from six horse breeds genotyped with 50 microsatellite markers using STRUCTURE analysis with eight inferred clusters.

Breed[1]	Number of individuals	Inferred clusters[2]							
		I	II	III	IV	V	VI	VII	VIII
AR	36	0.009	0.013	0.005	0.003	0.004	0.015	0.949	0.003
COM	33	0.009	0.003	0.006	0.011	0.006	0.005	0.008	0.952
FM	136	0.012	0.014	0.087	0.863	0.005	0.007	0.004	0.009
ICL	33	0.011	0.002	0.010	0.005	0.954	0.004	0.007	0.006
WB	50	0.317	0.126	0.009	0.008	0.006	0.505	0.023	0.007
TB	31	0.013	0.961	0.003	0.002	0.002	0.011	0.006	0.002

[1]AR, Arabian; COM, Comtois; FM, Franches-Montagnes; ICL, Icelandic Horse; WB, European Warmblood; TB, Thoroughbred.
[2]Numbers represent the proportion of individuals from each breed assigned to a given cluster.

Table 3 Assignment of 319 horses genotyped with 50 microsatellite loci to six pre-defined breeds using the algorithm of GeneClass2.

	Breed[1]					
	AR	COM	FM	ICL	WB	TB
Number of individuals	36	33	136	33	50	31
Assigned to the pre-defined population (%)	97	94	99	97	100	100
Assigned to the pre-defined population as well as another population (%)	78[2]	0	0	0	0	100[2]
Not assigned to the pre-defined but to at least another population (%)	3[2]	0	0	0	0	0
Neither assigned to the pre-defined nor to another population (%)	0	6	1	3	0	0

[1]AR, Arabian; COM, Comtois; FM, Franches-Montagnes; ICL, Icelandic Horse; WB, European Warmblood; TB, Thoroughbred.
[2]Assignments to another population were in all cases to WB only.

(14.8%). The average levels of observed and expected heterozygosity are also comparable with other studies (Behara *et al.* 1998; Bjørnstad *et al.* 2000; Cañon *et al.* 2000; Aberle *et al.* 2004). The smaller numbers of alleles in TB and AR confirm earlier findings (Bowling & Ruvinsky 2000; Tozaki *et al.* 2001, 2003) and could be a consequence of the higher degree of inbreeding because of the narrow genetic base of these breeds (Cunningham *et al.* 2001).

In FM, there was no relationship between the individual coefficient of inbreeding and the degree of homozygosity. This deviation from theory could be due to segregation distortion (Zöllner *et al.* 2004); to the fact that, from birth to three years of age, horses with a higher degree of homozygosity may have lower fitness; to questionable parentage of ancestors; or to mistyping of homozygous animals as heterozygous. Genotyping errors can be excluded because of extensive checking, and questionable parentage seems quite unlikely. A significant negative correlation of inbreeding coefficient and heterozygosity was found in German draught populations (Aberle *et al.* 2004). No significant association was observed between inbreeding, microsatellite heterozygosity and morphological traits in the Lipizzan horse (Curik *et al.* 2003). In cattle, small inbred populations do not always exhibit reduced allele numbers and/or low heterozygosity values (Maudet *et al.* 2002).

The different approaches used to measure genetic diversity between horse breeds all assign individuals to groups *a posteriori* using multilocus genotypes. The neighbour-joining phylogram constructed from allele-sharing distances formed clearly defined clusters for the ICL, COM, AR and FM breeds with no or only restricted introgression. Overlapping clustering could be seen for WB and TB, resulting from frequent grading-up of WB with TB. The Bayesian approach detects not only population structures by the identification of clusters, but also allows a probabilistic estimate of the proportion with which an individual belongs to one of the inferred clusters. These tools are now widely applied for genetic analysis (Beaumont & Rannala 2004; Pearse & Crandall 2004; Krüger *et al.* 2005). They are also in discussion for forensic applications (Manel *et al.* 2002; Bamshad *et al.* 2003; Jobling & Gill 2004). Eight inferred clusters fitted our data best. Horses of breeds that do not allow outbreeding (AR, COM, ICL and TB) were clearly unified in their own cluster. Only a few exceptions have a membership <0.8 to their own cluster. In contrast, WB horses are very heterogeneous, resulting from permanent migration of WB horses among the European countries and from introgression of TB (Riggenbach *et al.* 2005).

Despite the known introgression of WB into the FM breed, all FM in the neighbour-joining tree grouped together in the same branch. The Bayesian approach showed distinct allocation and revealed that the FM can be grouped into subpopulations that differ in the proportion of WB and in the composition of WB founder stallions.

The direct assignment test with the six breeds showed that the sampled COM, FM, ICL and WB have more suitable

multilocus genotype compositions for unequivocally assigning individuals to their native breed (94%, 99%, 97% and 100% respectively) than the AR and TB (19% and 0%). The latter two breeds had additional assignment probabilities in WB, which reflects the frequent grading-up of WB with TB and occasionally with AR. Therefore, it will be very difficult to develop a reliable test for AR and especially TB to solve the question if an AR or a TB horse can be assigned to its 'true' breed based on a multilocus genotype in the presence of WB. However, the WB horses were all assigned to the pre-defined population and none were co-assigned to AR or TB.

In this study, allele frequencies of 50 microsatellite loci differentiated populations that do not have significant interbreeding or introgression with other populations. Individual breed assignments, which require a high certainty of origin in forensic applications, are very difficult for breeds that are used for outcrosses. In such cases, other criteria, e.g. mtDNA typing results and breed-specific phenotypic traits, may be added to obtain more accurate conclusions. The findings of this study will support the conservation effort for Franches-Montagnes by assigning or excluding individuals from the native breed.

Acknowledgements

We thank Drs Cecilia Penedo and Teruaki Tozaki for helpful primer information. This work received financial support from the Swiss Federal Office of Agriculture. The authors thank the reviewers for constructive comments on the manuscript.

References

Aberle K.S., Hamann H., Drögemüller C. & Distl O. (2004) Genetic diversity in German draught horse breeds compared with a group of primitive, riding and wild horses by means of microsatellite DNA markers. *Animal Genetics* **35,** 270–7.

Bamshad M.J., Wooding S., Watkins W.S., Ostler C.T., Batzer M.A. & Jorde L.B. (2003) Human population genetic structure and inference of group membership. *American Journal of Human Genetics* **72,** 578–89.

Beaumont M.A. & Rannala B. (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics* **5,** 251–61.

Behara A.M.P., Colling D.T., Cothran E.G. & Gibson J.P. (1998) Genetic relationships between horse breeds based on microsatellite data: applications for livestock conservation. *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production* **28,** 119–22.

Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. (1996–2002) *GENETIX 4.04, Logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Bjørnstad G. & Røed K.H. (2002) Evaluation of factors affecting individual assignment precision using microsatellite data from horse breeds and simulated breed crosses. *Animal Genetics* **33,** 264–70.

Bjørnstad G., Gunby E. & Roed K.H. (2000) Genetic structure of Norwegian horse breeds. *Journal of Animal Breeding and Genetics* **117,** 307–17.

Bowcock A.M., Ruiz-Linares A., Tomfohrde J., Minch E., Kidd J. R. & Cavalli-Sforza L.L. (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature,* **368,** 455–7.

Bowling A.T. & Ruvinsky A. (2000) *The Genetics of the Horse*. CAB International, Oxon, UK. ISBN 0 85199 429 6.

Bruford M.W., Bradley D.G. & Luikart G. (2003) DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics* **4,** 900–10.

Cañon J., Checa M.L., Carleos C., Vega-Pla J.L., Vallejo M. & Dunner S. (2000) The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. *Animal Genetics* **31,** 39–48.

Cunningham E.P., Dooley J.J., Splan R.K. & Bradley D.G. (2001) Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. *Animal Genetics* **32,** 360–4.

Curik I., Zechner P., Sölkner J., Achmann R., Bodo I., Dovc P., Kavar T., Marti E. & Brem G. (2003) Inbreeding, microsatellite heterozygosity, and morphological traits in Lipizzan horses. *Journal of Heredity* **94,** 125–32.

Felsenstein J. (1993–2002) *Phylogeny Inference Package (PHYLIP)*. Genomes sciences, Department of Genetics, University of Washington, Seattle, WA. Software available at http://evolution.gs.washington.edu/phylip.html.

Jobling M.A. & Gill P. (2004) Encoded evidence: DNA in forensic analysis. *Nature Reviews Genetics* **5,** 739–52.

Krüger K., Gaillard C., Stranzinger G. & Rieder S. (2005) Phylogenetic analysis and species allocation of individual equids using microsatellite data. *Journal of Animal Breeding and Genetics* **122**(Suppl. 1), 78–86.

Manel S., Berthier P. & Luikart G. (2002) Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. *Conservation Biology* **16,** 650–9.

Maudet C., Luikart G. & Taberlet P. (2002) Genetic diversity and assignment tests among seven French cattle breeds based on microsatellite DNA analysis. *Journal of Animal Science* **80,** 942–50.

Minch E., Ruiz-Linares A., Goldstein D.B., Feldman M.W. & Cavalli-Sforza L.L. (1998) *Microsat2: A Computer Program for Calculating Various Statistics on Microsatellite Allele Data*. Department of Genetics, Stanford University, Stanford, CA. Software available at http://hpgl.stanford.edu/projects/microsat/).

Nei M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* **70,** 3321–3.

Nei M. & Chesser R.K. (1983) Estimation of fixation indices and gene diversities. *Annals of Human Genetics* **47,** 253–9.

Paetkau D., Slade R., Burden M. & Estoup A. (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13,** 55–65.

Pearse D.E. & Crandall K.A. (2004) Beyond F_{ST}: analysis of population genetic data for conservation. *Conservation Genetics* **5,** 585–602.

Piry S., Alapetite A., Cornuet J.-M., Paetkau D., Baudouin L., & Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95,** 536–9. Available at http://www.montpellier.inra.fr/CBGP/softwares/.

Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155,** 945–59. Software STRUCTURE available at: http://pritch.bsd.uchicago.edu/.

Raymond M. & Rousset F. (1995) GENEPOP (version 1.2) population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86,** 248–9.

Riggenbach Ch., Stranzinger G., Poncet P.A., Glowatzki M.L., Muntwyler J., Gaillard C. & Rieder S. (2005) Molekulargenetische Untersuchungen beim Einsiedler-Pferd (Molecular genetic studies on the "Einsiedler" horse population). *Schweizer Archiv für Tierheilkunde* **147,** 195–204.

Rosenberg N.A. (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* **4,** 137–8. http://www.cmb.usc.edu/~noahr/distruct.html.

Rosenberg N.A., Li L.M., Ward R. & Pritchard J.K. (2003) Informativeness of genetic markers for inference of ancestry. *American Journal of Human Genetics* **73,** 1402–22.

Saitou N. & Nei M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4,** 406–25.

Tozaki T., Kakoi H., Mashima S., Hirota K., Hasegawa T., Ishida N., Miura N., Choi-Miura N.H. & Tomita M. (2001) Population study and validation of paternity testing for Thoroughbred horses by 15 microsatellite loci. *Journal of Veterinary Medical Science* **63,** 1191–7.

Tozaki T., Takezaki N., Hasegawa T., Ishida N., Kurosawa M., Tomita M., Saitou N. & Mukoyama H. (2003) Microsatellite variation in Japanese and Asian horses and their phylogenetic relationship using a European horse outgroup. *Journal of Heredity* **94,** 374–80.

Zöllner S., Wen X., Hanchard N.A., Herbert M.A, Ober C. & Pritchard J.K. (2004) Evidence for extensive transmission distortion in the human genome. *American Journal of Human Genetics* **74,** 62–72.

Supplementary Material

The following supplementary material is available from http://www.blackwell-synergy.com:

Figure S1. Relationship between the coefficient of inbreeding of 107 FM horses and the degree of homozygosity based on 50 microsatellite loci.

Table S1. List of microsatellite loci amplified in four multiplex PCRs.

Table S2. Number of alleles and unbiased coefficient of gene differentiation (G_{ST}) for 50 microsatellites in six horse breeds (sorted by G_{ST}).

Table S3. Estimated assignment probabilities of 319 horses to six breeds (GeneClass2).