GENETIC DIVERSITY IN A POPULATION OF LUSITANO STALLIONS BORN OVER THREE DECADES

Sofia Abreu Ferreira[1], Carlos Ferreira[1], Maria Ramalho[1], Maria Gonçalves[1], Nuno Carolino[2], Maria Susana Lopes[3]

[1] Laboratório de Genética Molecular de Alter, Unidade Estratégica de Investigação e Serviços de Biotecnologia e Recursos Genéticos, Instituto National de Investigação Agrária e Veterinária (INIAV, IP), Coudelaria de Alter, 7440-152 Alter do Chão, Portugal.

[2] Estação Zootécnica Nacional, Unidade Estratégica de Investigação e Serviços de Biotecnologia e Recursos Genéticos, Instituto Nacional de Investigação Agrária e Veterinária (INIAV, IP), Portugal; CIISA – Faculdade de Medicina Veterinária da Universidade de Lisboa, Portugal; Sociedade Portuguesa de Recursos Genéticos Animais (SPREGA), Portugal; Escola Universitária Vasco da Gama, Coimbra, Portugal.

[3] Universidade dos Açores, Centro de Biotecnologia dos Açores, Rua Capitão João D'Ávila, 9700-042 Angra do Heroísmo, Portugal.

Address for correspondence:

Sofia Abreu Ferreira; Instituto National de Investigação Agrária e Veterinária (INIAV, IP), Unidade Estratégica de Investigação e Serviços de Biotecnologia e Recursos Genéticos, Laboratório de Genética Molecular de Alter, Coudelaria de Alter, 7440-152 Alter do Chão, Portugal; Phone: + 351-245610070; Fax: +351-245610071; E-mail: sofia.ferreira@iniav.pt

Maria Susana Lopes; Centro de Biotecnologia dos Açores, Universidade dos Açores Rua Capitão João D'Ávila, 9700-042 Angra do Heroísmo, Portugal; Phone: + 351-295-402-200; FAX: + 351-295-402-205; E-mail: maria.sb.lopes@uac.pt

1. INTRODUCTION

The Lusitano Horse (LH) is, besides the Garrano, the Sorraia and the Pónei da Terceira, economically, the most important Portuguese autochthonous equine breed and one of the most ancients in the word. The LH has its origins in wild and domesticated animals from the Iberian Peninsula intermixed with animals brought by several migration events that occurred in the past (Oom 1992; Lopes et al. 2005). The horses from this gene pool contributed to the development of many other modern European horses and were later introduced and dispersed throughout the Americas, where they were the founders of numerous New World breeds (Gonzaga 2004; Luís et al. 2006). According to the DAD-IS Database (2022), the actual population of LH is 5200 individuals registered in the studbook of which 1500 are breeding males and 3500 breeding females distributed over 500 studs.

The breed is reared in many countries, with Brazil recording the second largest population (da Silva Faria et al. 2018) followed by France and Spain (Vicente et al. 2012).

Sires enter reproduction at seven years and are trained and used in equestrian activities for several years, that implies that horses delay their long reproductive career of breeding individuals and leads to high generation interval, close to 11 years (Vicente et al. 2012) and consequently to the possibility of an increase in inbreeding (Cozzi et al. 2022). Thus, is very important to monitor the genetic diversity and the results could be compared with those of other breeds.

In this study we analyzed a group of 2699 stallions of the LH breed, born between 1985 and 2010 and registered in the Studbook ("Livro de Adultos", Book of Breeders section) providing information on the genetic variation and giving an indication of the levels of intra-breed variability. This group of stallions contains representatives of all possible paternal lineages of the Lusitano breed born over three decades.

2. MATERIALS AND METHODS

2.1 Population sampling

From the 41126 Lusitano samples stored at the Laboratório de Genética Molecular de Alter, between August 1998 and December 2012, we selected all males, alive (at the time of the study), not castrated, registered in the Book of Breeders and born between 1985 and 2010, making a total of 2699 animals. These animals were genotyped in a routine setting in paternity tests.

2.2 DNA extraction

Genomic DNA was extracted from whole blood samples, collected in 10 ml tubes with EDTA anti-coagulant, using the saline method (Montgomery and Sise 1990), resuspended in TE (pH 8.0) and preserved at 4 °C. The DNA was quantified by spectrophotometry and diluted to a final concentration of 3 ng/µl.

2.3 PCR amplification and Genotyping

The laboratory used six STRs (HMS3, HMS7, HTG4, HTG10, VHL20 and ASB2) until 2001, since they are extremely polymorphic, especially in breeds with Iberian influence (Luís et al. 2002). In 2001, three additional markers (AHT4, AHT5 and HMS6) were added from a panel recommended by the International Society for Animal Genetics (ISAG), making a total of nine markers that were systematically used until 2009.

From 2009 on, routinely the laboratory used 17 markers, of which 12 belong to the core panel recommended by ISAG (AHT4, AHT5, ASB2, HMS2, HMS3, HMS6, HMS7, HTG10, HTG4, VHL20, ASB17 and ASB23) and 5 to the additional ISAG STR panel (HTG6, HTG7, UCDEQ425, HMS1 and LEX3) (Table 2). From the additional panel of *loci* analyzed, LEX3 was excluded, since this marker is heterosomic, connected to the X-chromosome, meaning that it is homozygous in males, causing significant deviations to the Hardy-Weinberg Equilibrium (HWE) and high differences between expected and observed heterozygosity values (Zabek and Fornal 2009).

2.4 For 6 and 9 STRs

PCR amplification was performed with a 10 µL final solution containing 1 µL of genomic DNA, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, fluorescent and non-fluorescent primers with a concentration ranging from 0.5 pM to 3 pM, 0.5 U of Taq DNA polymerase and 1X reaction buffer. The amplification of the fragments was performed in a MJ-Research PTC-100TM thermal cycler (MJ Research Inc, Watertown. MA, USA), initiated with a denaturation step at 95 ºC (5 min), followed by 5 cycles with denaturation at 95 ºC (30 sec), hybridization at 58 ºC (30 sec) and extension at 72 ºC (30 sec) and 35 cycles with denaturation at 95 ºC (30 sec), hybridization at 57 ºC (30 sec) and extension at 71 ºC (30 sec) the last extension was extended for 10 min at 72 ºC. The *loci* HMS7, HMS3, HTG4 and VHL20, were amplified in tetraplex, the *loci* ASB2, HMS6 and HTG10 in triplex and the *loci* AHT4 and AHT5, in singleplex. PCR products were separated on denaturing 6 % polyacrylamide gels, 0.25 mm wide, using LI-COR 4000 and 4200 automated fluorescence sequencers (Li-Cor Biosciences, Lincoln, NE, USA). The resulting images were analysed with RFLPscan v.3.1 software (Scanalytics, CPS Inc., USA). The size of the DNA fragments was identified based on the 50-350 bp Licor internal size standard.

2.5 For 17 STRs

PCR amplification was performed using a commercial kit from Applied Biosystems (Applera Europe, Germany) with 17 primer pairs from which the forward primers are fluorescently labelled, following the manufacturer's instructions. Amplification was carried out in a MJ-Research PTC-100 ™ thermal cycler (MJ Research Inc., Watertown, MA, USA), starting with 95 0C denaturation for 10 min followed by 30 cycles with denaturation at 95 0C (30 sec), hybridization at 60 0C (30 sec) and extension at 72 0C (30 sec) with a last extension of 60 min at 72 0C, finishing at 4 0C. The amplified products were separated in an ABI 3130 Capillary Sequencer, with detection capacity for 5 fluorescent markers, 6-FAM, VIC, NED and PET. Each sample was run with the LIZ 500 marker and the G5 filter. The sizes of the resulting fragments for the 17 microsatellites were analyzed with Genemapper 4.0. This method was regularly checked and validated by the ISAG horse parentage comparison tests.

2.6 Statistical analysis

The analysis of the genetic diversity of the 2699 breeding males was estimated using the CERVUS 3.0.7 program (Marshall et al. 1998) by determining the total and mean number of alleles, their frequencies of observed (Ho) and of expected heterozygosity (He) (Nei 1978), the polymorphic information content (PIC) (Botstein et al. 1980) and the estimated frequency for the occurrence of null alleles (Summers and Amos 1997). Deviations from HWE with Fisher's exact test (Guo and Thompson 1992) were calculated by using Genepop 3.4 (Raymond and Rousset 1995). This test indicates the probability of error in rejecting the null hypothesis (H0) that if the population is in HWE (Raymond and Rousset 1995). The deficit of heterozygosity *per locus* (F_{IS}) and the mean value for all *loci* were estimated according to Weir and Cockerham (1984).

The Bayesian approach of STRUCTURE (Pritchard et al. 2000) was used to infer the genetic structure of the population. A 20.000 initial burn-in was used to minimize the effect of the starting configurations, followed by 100.000 MC iterations, as recommended by Falush et al. (2007) with 10 independent replicates each. Several sets of inferred clusters where tested to determine the most appropriate number of clusters for modeling the data. All runs used an admixture model with correlated frequencies and the parameter of individual admixture alpha set to be the same for all clusters and with a uniform prior.

The most likely number of clusters (K) was determined by calculating ΔK according to the Evanno method (Evanno et al. 2005) in Structure Harvester (Earl and von Holdt 2012).

3. RESULTS

For For the total of 2699 male Lusitano animals 131 alleles were obtained for the 16 *loci* analyzed, with a mean number of alleles of 8.19±2.40 *per locus*. The number of alleles *per locus* varied between 3 (HTG7) and 13 (ASB17), Ho ranged from 0.191 (HTG7) to 0.819 (VHL20), and He varied between 0.207 (HTG7) and 0.828 (ASB17)

(**Table 1**). These values indicate that for the group of animals analyzed, the least polymorphic *locus* is HTG7 and the most polymorphic one is ASB17. The observed heterozygosity was lower than the expected one for all *loci*, with mean values of 0.678±0.151 and 0.700±0.155, respectively, except for UCDEQ425.

Locus	N	Fragments Size (bp)	n	HO	HE	PIC
AHT4 [C]	2157	146-162	9	0.758	0.789	0.762
AHT5 [C]	2100	128-146	8	0.759	0.792	0.759
ASB2 [C]	2675	222-254	12	0.774	0.788	0.762
HMS2 [C]	844	218-238	6	0.743	0.768	0.731
HMS3 [C]	2689	146-168	8	0.793	0.824	0.803
HMS6 [C]	2168	159-169	7	0.597	0.606	0.570
HMS7 [C]	2688	173-187	8	0.657	0.697	0.644
HTG10 [C]	2689	91-111	10	0.691	0.722	0.695
HTG7 [A]	2697	129-141	7	0.642	0.648	0.592
HTG4 [C]	2692	89-107	10	0.819	0.827	0.803
VHL20 [C]	761	97-123	13	0.774	0.828	0.809
ASB17 [C]	900	184-212	8	0.781	0.797	0.765
ASB23 [C]	993	82-100	6	0.656	0.682	0.625
HTG6 [A]	893	120-128	3	0.191	0.207	0.194
UCDEQ425 [A]	713	224-244	9	0.658	0.656	0.594
HMS1 [A]	847	176-192	7	0.551	0.571	0.499
Mean ± SE	-	-	8.19 ±2.40	0.678±0.151	0.700±0.155	0.678±0.158

Table 1. Genetic variability of the analyzed population for 16 *loci*.

Source: No. of alleles per *locus* (n); number of individuals analyzed (N); Heterozygosity observed (HO); Heterozygosity expected (HE); Polymorphic information content (PIC); ([C]) Markers from the ISAG core STR panel; ([A]) Marker from the ISAG additional STR panel.

All *loci* are very informative, with the PIC higher than 0.5, except for HTG7. Values ranged from 0.194 (HTG7) to 0.809 (ASB17).

According to the HWE test *per locus* with P <0.05, only 4 *loci* (HTG4, VHL20, HTG6 and HMS1) are in HWE (**Table 2**). All *loci* present deficit in heterozygotes, except for ASB2 and UCDEQ425, which corroborates the values obtained for the observed heterozygosity. For the total *loci* we observed a significant deficit of heterozygotes (F_{IS}=0.034), as confirmed by the test of χ^2, indicating that the population is not in equilibrium.

Locus	P	F_{IS}
AHT4 [C]	0.0014	0.0636
AHT5 [C]	0.0004	0.0347
ASB2 [C]	0.0433	-0.0121
HMS2 [C]	0.0375	0.0256
HMS3 [C]	0.0005	0.0356
HMS6 [C]	0.0119	0.0353
HMS7 [C]	0.0000	0.0382
HTG10 [C]	0.0031	0.0561
HTG7 [A]	0.0236	0.0629
HTG4 [C]	0.2324	0.0369
VHL20 [C]	0.3516	0.0402
ASB17 [C]	0.0000	0.0636
ASB23 [C]	0.0398	0.0084
HTG6 [A]	0.1658	0.0401
UCDEQ425 [A]	0.0137	-0.0208
HMS1 [A]	0.0555	0.0240
Mean ± SE	0.0612±0.1035	0.0340±0.0241

Table 2.

Source: Values obtained for 16 *loci* for the Hardy-Weinberg Equilibrium test (Genepop 3.4) with P <0.05; Probability (P); Heterozygosity Deficit (F_{IS}).

The Bayesian analysis carried out in STRUCTURE revealed that the most likely number of clusters was two, as ΔK showed a substantial increase at K=2 when K varied from 1 to 15, although some substructure was observed at K=5 and K=9 (Fig. 1).

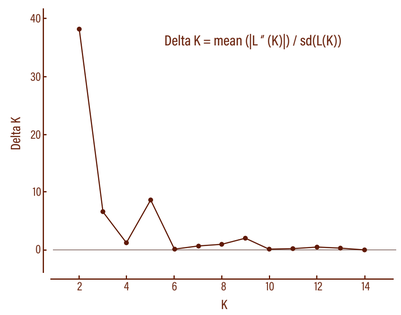

Fig. 1.

Source: Plot of ΔK obtained with STRUCTURE HARVESTER for detecting the number of genetic groups that best fit the Lusitano individuals analyzed.

4. CONCLUSIONS

The genetic diversity detected shows that it is similar to other equine breeds reported in the literature. However, the breeds' diversity should be monitored and taken into account in order to guide decisions on the management and conservation of the paternal lineages of the Lusitano horse.

F_{is} value indicated that overall Lusitano presents reasonable levels of heterozygotes and suggesting that inbreeding is considerately controlled.

REFERENCES

1. Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length polymorphisms. Am. J. Hum. Genet. 32, 314-331. 0002-9297/80/3203-0013$01.58
2. Cañon, J., Checa, M.L., Carleos, C., Veja-Pla, J.L., Vallejo, M., Dunner, S., 2000. The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. Anim. Genet. 31, 39-48. 10.1046/j.1365-2052.2000.00591.x
3. Costa Ferreira, J., Oom, M.M., 1989. Consanguinidade na raça Lusitana. Revista Portuguesa de Ciências Veterinárias, Vol. LXXXIV, 490, 77-89.
4. Cozzi, M.C., Valiati, P., Longeri, M., Ferreira, C., Abreu Ferreira, S. 2022 Genetic Variability Trend of Lusitano Horse Breed Reared in Italy. Animals 12, 98. 10.3390/ani12010098
5. DAD-IS (Domestic Animal Genetic Diversity Information System), 2022. https://www.fao.org/dad-is/en/ (accessed 20 February 2022).
6. da Silva Faria, R.A., Andrade Vicente, A.P., Duarte Guedes dos Santis, R.I., Marchi Maiorano, A., Curi, R.A., Loyola Chardulo, L.A., Vascovelos Silva, J.A. 2018. Genetic diversity of Lusitano Horse in Brazil using pedigree information. J Equine Vet Sci 69, 149-158. 10.1016/j.jevs.2018.07.009
7. Earl, D.A., von Holdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resour 4(2), 359-361. 10.1007/s12686-011-9548-7 Core version: vA.2 July 2014; Plot version: vA.1 November 2012; Web version: v0.6.94 July 2014.
8. Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611-2620. 10.1111/j.1365-294X.2005.02553.x
9. Falush, D., Stephens, M., Pritchard, J.K., 2007 Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7, 574–8. 10.1111/j.1471-8286.2007.01758.x
10. Felicietti, M., Lopes, M.S., Verini-Supplizi, A., da Câmara Machado, A., Silvestrelli, M., Mendonça, D., Disti, O., 2010. Genetic diversity in the Maremmano horse and relationships with other European breeds. Anim. Genet. 41, 53-55. 10.1111/j.1365-2052.2010.02102.x
11. Gonzaga, P.G., 2004. A History of the Horse: The Iberian Horse From Ice Age to Antiquity. JA Allen and Co. Ltd. London.
12. Guo, S.W., Thompson, E.A., 1992. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics 48, 361-372. 10.2307/2532296
13. Lopes, M.S., Abreu Ferreira, S., Ferreira, C., da Costa Ferreira, J., Mendonça, D., da Câmara Machado, A., 2015. Evolution of the genetic diversity of the Lusitano breed over the past 4 decades. 11th Dorothy Russell Havemeyer International Equine Genome Mapping Workshop, July 22th-25th, Hannover, Germany.
14. Lopes, M.S., 2011. Molecular Tools for the Characterisation of the Lusitano Horse. Tese de Doutoramento, Universidade dos Açores.
15. Lopes, M.S., Mendonça, D., Cymbron, T., Valera, M., Costa-Ferreira, J., da Câmara Machado, A., 2005. The Lusitano horse maternal lineage based on mitochondrial D-loop sequence variation. Anim. Gen. 36, 196–202. 10.1111/j.1365-2052.2005.01279.x
16. Luís, C., Gothran, E.G., Oom, M.M., 2002. Microsatellites in Portuguese autochthonous horse breeds: usefulness for parentage testing. Gen. Mol. Biol. 25, 131-134. 10.1590/S1415-47572002000200003
17. Luís, C., Bastos-Silveira, C., Cothran, E.G., Oom, M.M., 2006. Iberian origins of New World horse breeds. J. Hered. 97, 107-113. 10.1093/jhered/esj020
18. Luís, C., Juras, R., Oom, M., Cothran, E., 2007. Genetic diversity and relationships of Portuguese and other horse breeds based on protein and microsatellite loci variation. Anim. Genet. 38, 20–27. 10.1111/j.1365-2052.2006.01545.x
19. Marshall, T.C., Slate, J., Kruuk, E.B., Pemberton, M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7, 639-655. 10.1046/j.1365-294x.1998.00374.x
20. Montgomery, G.W., Sise, J.A., 1990. Extraction of DNA from sheep white blood cells. New Zealand J. Agric. Res. 33,437-441. 10.1080/00288233.1990.10428440
21. Nei, M., Li, W., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. USA 76, 5269-5273. 10.1073/pnas.76.10.5269
22. Oom, M.M., 1992. O Cavalo Lusitano. Uma Raça em Recuperação. Tese de Doutoramento, Universidade de Lisboa.
23. Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155(2), 945-959. 10.1534/genetics.116.195164
24. Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): Population Genetics software for Exact Tests and Ecumenicism. J. Hered. 86, 248-249. 10.1093/oxfordjournals.jhered.a111573
25. Solis, A., Jugo, B.M., Meiaux, J.C., Iriondo, M., Mazon, A.I., Aguirre, A.I., Vicario, A., Estomba, A., 2005. Genetic Diversity Within and Among Four South European Native Horse Breeds Based on Microsatellite DNA Analysis: Implications for Conservation. J. Hered. 96, 670-678. 10.1093/jhered/esi123
26. Summers, K., Amos, W., 1997 Behavioral, ecological and molecular genetics analyses of reproductive strategies in the Amazonian dart-poison frog, Dendrobates ventrimaculatus. Behav Ecol. 8, 260-267. 10.1093/beheco/8.3.260
27. van de Goor, L.H., van Haeringen, W.A., Lenstra, J.A., 2011. Population studies of 17 equine STR for forensic and phylogenetic analysis. Anim Genet. 42(6), 627-633. 10.1111/j.1365-2052.2011.02194.x
28. Vdovina, N.V., Yuryeva, I.B., 2021. Monitoring for the genetic structure of Mezen breed of horses in terms of DNA microsatellites. Vavilovskii Zhurnal Genet Selektsii;;25(2),202-207. 10.18699/VJ21.024
29. Vicente, A.A., Carolino, N.; Gama, L.T., 2012. Genetic diversity in the Lusitano horse breed assessed by pedigree analysis. Livest. Sci. 148, 16-25. 10.1016/j.livsci.2012.05.002
30. Vicente, A.A., 2015. Characterization and selection of the Lusitano horse breed. PhD Thesis, Lisbon University
31. Weir, B., Cockerham, C., 1984. Estimating F-statistics for the analysis of population structure. Evol. 38, 1358-1370. 10.2307/2408641
32. Zabek, T., Fornal, A., 2009. Evaluation of the 17-Plex STR KIT for Parentage Testing of Polish Coldblood and Hucul Horses. Ann. Ani. Sci. 4, 363-372
33. Zuccaro, A., Bordonaro, S., Criscione, A., Guastella, A.M., Perrotta, G., Blasi, M., D'Urso, G., Marletta, D., 2008. Genetic diversity and admixture analysis of Sanfratellano and three other Italian horse breeds assessed by microsatellite markers. Animal 2(7), 991-998. 10.1017/S1751731108002255.