GENETIC DIVERSITY AND POPULATION STRUCTURE IN REMNANT SUBPOPULATIONS OF NORDESTINO HORSE BREED

DIVERSIDADE GENÉTICA E ESTRUTURA POPULACIONAL DE SUBPOPULAÇÕES REMANESCENTES DA RAÇA EQUINA NORDESTINO

Pires, D.A.F.[1]*; Coelho, E.G.A.[2]; Melo, J.B.[3]; Oliveira, D.A.A.[2]; Ribeiro, M.N.[1]; Gus Cothran, E.[4]; Juras, R.[4] and Khanshour, A.[4]

[1]Departamento de Zootecnia. Universidade Federal Rural de Pernambuco. Recife-PE. Brasil. *dna@zootecnista.com.br

[2]Laboratório de Genética da Escola de Veterinária. Universidade Federal de Minas Gerais. Belo Horizonte-MG. Brasil.

[3]Departamento de Estudos Básicos e Instrumentais. Universidade Estadual do Sudoeste da Bahia. Itapetinga-BA. Brasil.

[4]Department of Veterinary Integrative Biosciences. Texas A & M University. College Station, TX. USA.

ADDITIONAL KEYWORDS

Allelic diversity. Bottleneck effect. Genetic structure. Microsatellite.

PALAVRAS CHAVE ADICIONAIS

Diversidade alélica. Efeito de gargalo genético. Estrutura genética. Microssatélite.

SUMMARY

This study analyzed four remnant subpopulations of Nordestino horse breed to detect genetic structure and diversity through 14 microsatellite markers. Hair root follicles from a total of 393 horses were collected. There were 61 animals from Salitre Valley (JUAZ-BA) located at Bahia state, 89 from North and Central North ecoregions located at Piauí state (NCEN-PI), 185 animals from Sertão and Sertão do São Francisco ecoregions (SERT-PE) and 58 animals from Agrestina city (AGRE-PE) located at Pernambuco state. Genetic diversity, genetic differentiation and bottleneck effects were examined in the 4 remnant subpopulations of Nordestino horse breed. There was high allelic diversity and the F_{is} value did not show evidence of a significant predominance of mating among relatives, probably because of crossbreeding among populations. Recent bottleneck effects were not detected in the 4 subpopulations, but the IAM and TPM model did suggest a bottleneck effect. This may be a reflection of the decreased number of breeding animals caused by castration of males, mechanization processes and changes in life style in the rural areas. The bottleneck event was not enough to lead a genetic differentiation among the 4 remnant subpopulations of Nordestino horse. There was no evidence of genetic differentiation, so the 4 subpopulations formed one genetic group.

RESUMO

O presente trabalho teve como objetivo avaliar a estrutura e diversidade genética de quatro subpopulações do cavalo Nordestino utilizando 14 marcadores microssatélites. Amostras de bulbo capilar de 393 cavalos foram coletadas, distribuídos em: 61 animais do vale do Salitre (JUAZ-BA) localizado no estado da Bahia, 89 das mesorregiões Norte e Centro-Norte do estado do Piauí, 58 animais da cidade de Agrestina (AGRE-PE) e 185 das mesorregiões Sertão e Sertão do São Francisco localizado no estado de Pernambuco (SERT-PE). A diversidade genética, diferenciação genética e efeito de gargalo genético foram avaliados nas 4 subpopulações do cavalo Nordestino. Observou-se elevada diversidade genética e valores de F_{is} que não evidenciaram níveis de consanguinidade significativa, provavelmente isso se deve à acasalamentos entre as subpopulações. Efeito de gargalo genético recente não foi detectado para as quatro subpopulações,

Recibido: 12-11-13. Aceptado: 20-3-14. *Arch. Zootec. 63 (242): 349-358. 2014.*

porém os modelos IAM e TPM sugeriram um efeito de gargalo genético não recente. Isso pode ser reflexo da redução no número de reprodutores decorrente da castração de machos, processo de mecanização e mudanças no estilo de vida das populações da zona rural. Eventos mais antigos do efeito de gargalo genético não foram suficientes para evidenciar uma diferenciação genética entre as 4 subpopulações, portanto elas formam um só grupo genético.

INTRODUCTION

The Nordestino horse is a local breed from the Brazilian Northeast that has adapted to semi-arid conditions of the Caatinga biome. The Nordestino horse breed is characterized by small animals, small head, strong hooves with deep frog, dark skin and good work ability in the semi-arid Northeast region (ABCCN, 1987; Melo, 2011). High insolation, few clouds, low forage available, rain shortage over the year and rocked soils are Northeast semiarid condition at Brazil, and the Nordestino horse can survive and keep its healthy status very well furthermore it can ride on rocked soils naturally without compromise its hooves because they are adapted. Other horse breeds could not survive under those conditions or they would have a low performance with degeneration of their characteristics. Molecular Genetic information's about Nordestino Horse does not exist until 2013.

The first breed association of the Nordestino horse was founded on 1974 and the headquarters was located at Recife city. However, in the 1990s the association was shut down until 2011 (Melo, 2011). As a result, the official record of the Nordestino horse stopped. Most of the registered animals were sold for slaughter. Nowadays, Nordestino horse lives without a registry and with a minimum or almost no management. The remaining Nordestino horses are at risk as a result of shutting the association down, increased crossbreeding that might lead as mischaracterization of the Nordestino horse in the future and the high percentage of castrated stallions (Melo, 2011). Add to this the mechanization of agricultural processes and the increase of automobiles for transportation in rural area as a threat for the Nordestino horse. Nordestino horse is in a different situation from other Brazilian breeds: they are horses of poor farmers in Brazil, there is no interest from Brazilian government to conserve or preserve them and nearly all people still have Nordestino horse are unaware the importance to preserve it as the unique horse able to survive with a good performance in Semiarid condition from Brazil.

One of the first steps in a conservation program for a local breed is the genetic characterization, It provides important information for conservation management in a country's national strategy for animal genetic resources (FAO, 2012). The aims of this study were characterized genetically four subpopulations of Nordestino horses from different regions of the Brazilian Northeast and to determine if those represent a single population of the Nordestino horses through microsatellite markers.

MATERIAL AND METHODS

SAMPLE COLLECTION AND LOCATION SITES

Horse hair roots were collected from São Francisco Valley ecoregion in the Bahia state (number of animals= 61), cities from North and Central-North ecoregions in the Piauí state (89), Agrestina city in the Pernambuco state (58), cities from Sertão do São Francisco and Sertão ecoregions in the Pernambuco state (185). All ecoregions are inside the Caatinga biome (climate is semiarid). The animals were split into 4 subpopulations that were named as: JUAZ-BA (n=61), NCEN-PI (n=89), the animals from Pernambuco state were divided up into 2 subpopulations AGRE-PE (n=58) and SERT-PE (n=185). All animals sampled were considered as typical of the remainder of Nordestino horse breed because they had phenotypic traits similar to last breed stan-

dard: animals with height at withers of male 138 cm (±8) and female 135 cm (±8), small head, strong hooves with deep frog and dark skin (ABCCN, 1987). A total of 137 samples of the Arabian horse breed from Brazilian studs were included as an outgroup using data from LGEV/UFMG (Laboratory of Genetic in the Animal Science Department at Veterinary School of Federal University of Minas Gerais-Brazil).

Extraction of genomic DNA, PCR, electrophoretic run and reading of electropherogram

DNA extraction was made according to Coelho *et al*. (2004). PCR solution was prepared with 5.0 µL of Phusion Flash Master Mix Enzyme-Finnzymes, 1.0 µL of ultra pure water, 3.0 µL of primer mix and 1.0 µL of DNA extracted. The panel for genotyping consisted of 14 microsatellites: AHT4, AHT5, ASB2, ASB17, HTG4, HTG6, HMS3, HMS6, HMS7, ASB23, HTG7, HTG10, LEX 33 and VHL20. Those microsatellite marks were recommended by FAO/ISAG (Hoffmann *et al.*, 2004; FAO/ISAG, 2011) to study genetic characterization in horse breeds around the world. The standardized was according to FAO/ISAG (2011). The specific set of microsatellite is used in paternity tests because of its highly polymorphic. There were 3 multiplex panels for PCR: one with annealing temperature of 60 °C (AHT4, AHT5, ASB17, ASB23, HMS6, HMS7, HTG4 and VHL20), other with 56 °C (ASB2, HMS3 and HTG10) and the third with 60 °C (LEX33, HTG6 and HTG7). The same annealing temperature in 2 multiplex panels was used due to the typing of LEX33, HTG6 and HTG7 which are not part of routine testing. For microsatellite amplification in the 3 multiplex panel sets: 98 °C for 10 s for activation step, followed by 34 cycles of 95 °C for 45 s (denaturation step), 56 °C or 60 °C (depending on multiplex panel) for 30 s (annealing) and 72 °C for 30 s (extension), and a final extension step of 60 min at 72 °C. Capillary electrophoretic run was performed by 0.3 µL of LIZ™ (standard molecular weight), 8.7 µL of Formamide Hi-Di (both products are Applied Biosystems), and 1 µL of mixture of the panels from 3 PCR products, per sample. PCR products were analyzed by using the ABI3130 of the applied biosystems for capillary electrophoresis run. Fragments sizes were determined with GeneMapper v.4.0 software of the Applied Biosystems. All 530 samples were genotyped at LGEV/UFMG, Minas Gerais- Brazil.

Statistical analysis

Genetic diversity within each of the 4 subpopulations was measured as the number of alleles per locus (*Na*), effective number of alleles per locus (*Ne*), observed (*Ho*) and unbiased expected (*UHe*) heterozygosities estimated per locus that were calculated using the GenAlex 6.4 (Peakall and Smouse, 2006). Allelic richness (*Ar*) per locus was calculated with FSTAT v.2.9.3.2 (Goudet, 1995). It was used a rarefied sample size of 58 diploids individuals per subpopulation to calculate *Ar*. Heterozygote deficit and deviations from Hardy-Weinberg Equilibrium (HWE) were estimated by GenePop v.4.1.1 (Rousset, 2008). Deviations from HWE were performed with Markov Chain parameters: 10 000 dememorization, 20 batches and 5000 iterations per batch. Polymorphic Information Content (PIC) was estimated using the Cervus v.3.0.3 software (Kalinowski *et al*., 2007).

Wright's (1951) F-statistics (F_{st}, F_{it} and F_{is}) was performed with 1000 bootstrap on confidence interval of 95 % and Gene differentiation coefficient (G_{st}) were calculated using GENETIX v.4.04 (Belkhir *et al*., 2003). Analysis of molecular variance (AMOVA) with permutations set to 999 was calculated with the GenAlex 6.4 (Peakall and Smouse, 2006) to study genetic differentiation among subpopulations (Φ_{pt} and F_{st} pairwise) and fractionate the genetic variance within and among subpopulation.

A Bayesian method to cluster the animals was used to analyze genetic structure of the

4 subpopulations with STRUCTURE 2.3.3 (Pritchard *et al*., 2000; Falush *et al*., 2003). Model admixture and correlated allele frequencies were used as the basis of which an individual may have mixed ancestry. The model estimates *a posteriori* probability given that an individual originated from subpopulations among the inferred groups. The K values (number of clusters) ranging from 1 to 10 and for each cluster 20 independent runs were repeated. It was used a burn-in period of 20 000 and 100 000 Monte Carlo Markov Chain (MCMC) repetitions in all replications in order to obtain the number of suitable groups at the end, as recommended by Falush *et al*. (2007). The four subpopulations were analysed along with the Arabian group. The results of the natural logarithm (Ln) of the probability of the data (Ln Pr (X/K)) were used for detection of the best value of K (Evanno *et al.*, 2005) and to identify which approach is most appropriate. After that, the replicates were summarized for each K using CLUMPP software (Jakobsson and Rosenberg, 2007). Graphical display was generated using the DISTRUCT v.1.1 (Rosenberg, 2004) and GhostView v.4.9 (Lang, 2006) programs. Genotyping data from Arab breed horse was used as an outgroup to Nei's D_A genetic distance and Structure software.

The BOTTLENECK v.1.2.02 (Cornuet and Luikart, 1996) program was used to detect possible bottleneck event on the 4 subpopulations of Nordestino horse breed. Three models of microsatellite evolution were evaluated (IAM-Infinite allele model; SMM-stepwise mutation model and TPM-two-phase model of mutation) with two tests: Sign test and Wilcoxon sign-rank test. The probability distribution was established using 1000 simulations. The Mode Shift Indicator that based on the shape of the allele frequency distribution was also performed. The values of average heterozygosity (*He*) and their probabilities (*H*) in the Sign test, under three models of microsatellite evolution (IAM, SMM and TPM) were calculated and used to measure the expected number of loci with heterozygosity excess.

RESULTS

GENETIC DIVERSITY

The 14 microsatellites markers were amplified in the 4 remnant populations of the Nordestino horse breed. A total of 115, 110, 108 and 123 alleles were found and the average number of alleles per locus was 8.214, 7.857, 7.714 and 8.786 for AGRE-PE, JUAZ-BA, NCEN-PI and SERT-PE, respectively **(table I)**. The highest effective number of alleles per locus by subpopulations for AGRE-PE was 7.097 (VHL20), JUAZ-BA was 7.390 (VHL20), NCEN-PI was 7.120 (HTG10) and SERT-PE was 7.090 (HTG10). Average allelic richness per subpopulation was higher than 7 **(table I)**. The observed heterozygosity (*Ho*) ranged from 0.517 (HTG6 at NCEN-PI) to 0.902 (VHL20 at JUAZ-BA). The unbiased expected heterozygosity (*UHe*) ranged from 0.530 (HTG6 at NCEN-PI) to 0.872 (VHL20 at JUAZ-BA). Ott (1992) reported a locus is considered highly polymorphic when heterozygosity is greater than 0.7 and almost all loci were equal or higher than 0.7. There was high allelic variation in the remnant subpopulations of Nordestino horse breed. Significant ($p<0.05$) deviation from HWE was detected for HTG6 and LEX33 loci at AGRE-PE, HMS3 at NCEN-PI and HMS6 at SERT-PE. However, there were not significant deviations from HWE when all microsatellite markers were analyzed together in their respective subpopulation. Heterozygote deficit was identified at HMS3 and ASB2 loci at SERT-PE, but in multiloci analyzes in their respective subpopulation there was no heterozygote deficit. It was not observed significant F_{is} values for both loci inside subpopulations and average for each subpopulation **(table I)** were not significant. There were high PIC values per locus ($p>0.555$); the exceptions were observed in the HTG6 and HTG7 loci at NCEN-PI

and HTG7 at SERT-PE that showed medium polymorphism information content.

GENETIC DIFFERENTIATION

Low genetic differentiation among the four remnant subpopulations of Nordestino horse breed was found. The F_{is}, F_{it} and F_{st} (Weir and Cockerham, 1984) average values in the Nordestino subpopulations were 0.012, 0.017 and 0.005, respectively and not

Table I. *Genetic diversity in remnant population of Nordestino horse breed.* (Diversidade genética em subpopulações remanescentes do cavalo Nordestino).

Sp/locus	Na	Ne	A r	Ho	UHe	HWE	F_{is}	PIC
AGRE-PE (n=58)								
AHT4	8	5.203	7.951	0.845	0.815	ns	-0.037^{ns}	0.780
AHT5	8	4.748	6.902	0.828	0.796	ns	-0.040^{ns}	0.759
ASB17	12	5.640	12.899	0.81	0.830	ns	0.024^{ns}	0.802
ASB2	10	5.980	8.951	0.897	0.840	ns	-0.068^{ns}	0.813
ASB23	9	6.943	8.902	0.897	0.863	ns	-0.039^{ns}	0.839
HMS3	7	5.524	7.000	0.759	0.826	ns	0.082^{ns}	0.795
HMS6	7	3.216	6.000	0.69	0.695	ns	0.008^{ns}	0.660
HMS7	7	4.229	5.998	0.776	0.770	ns	-0.007^{ns}	0.726
HTG10	10	6.323	10.000	0.776	0.849	ns	0.087^{ns}	0.823
HTG4	7	3.539	6.000	0.672	0.724	ns	0.071^{ns}	0.675
HTG6	6	2.628	6.949	0.603	0.625	*	0.035^{ns}	0.575
HTG7	5	2.600	4.951	0.586	0.621	ns	0.056^{ns}	0.574
LEX33	8	4.242	6.951	0.724	0.771	*	0.061^{ns}	0.731
VHL20	11	7.097	9.951	0.845	0.867	ns	0.025^{ns}	0.843
Average	8.214	4.851	7.815	0.765	0.778	ns	0.017^{ns}	-
SE	0.526	0.401	0.568	0.026	0.022	-	-	-
JUAZ-BA (n=61)								
AHT4	8	4.689	7.647	0.869	0.793	ns	-0.096^{ns}	0.759
AHT5	7	4.128	6.638	0.689	0.764	ns	0.100^{ns}	0.718
ASB17	13	4.027	13.434	0.754	0.758	ns	0.005^{ns}	0.727
ASB2	9	4.634	8.610	0.82	0.791	ns	-0.037^{ns}	0.759
ASB23	9	5.219	8.262	0.836	0.815	ns	-0.026^{ns}	0.783
HMS3	7	4.704	6.959	0.672	0.794	ns	0.154^{ns}	0.760
HMS6	6	3.015	6.000	0.672	0.674	ns	0.003^{ns}	0.640
HMS7	6	3.494	5.986	0.77	0.720	ns	-0.071^{ns}	0.665
HTG10	10	6.006	8.995	0.803	0.840	ns	0.044^{ns}	0.814
HTG4	6	3.822	6.000	0.787	0.744	ns	-0.057^{ns}	0.697
HTG6	7	2.467	5.647	0.525	0.600	ns	0.126^{ns}	0.560
HTG7	5	2.674	4.000	0.639	0.631	ns	-0.013^{ns}	0.579
LEX33	7	4.103	7.878	0.803	0.762	ns	-0.054^{ns}	0.717
VHL20	10	7.390	8.650	0.902	0.872	ns	-0.034^{ns}	0.850
Average	7.857	4.312	7.479	0.753	0.754	ns	0.002^{ns}	-
SE	0.573	0.351	0.594	0.027	0.020	-	-	-

SP= subpopulation; n= sample size; Na= number of alleles per locus; Ne= effective number of alleles; Ar= allelic richness; Ho and UHe= observed and unbiased expected heterozygosities; HWE= Hardy-Weinberg equilibrium; F_{is}= inbreeding coefficient value; PIC= polymorphism information content; SE= standard error; ns= no significant ($p<0.05$).

Table I (continuation). *Genetic diversity in remnant population of Nordestino horse breed.* (Diversidade genética em subpopulações remanescentes do cavalo Nordestino).

Sp/locus	Na	Ne	A r	Ho	UHe	HWE	F_{is}	PIC
NCEN-PI (n=89)								
AHT4	8	4.580	7.976	0.82	0.786	ns	-0.044^{ns}	0.751
AHT5	7	4.487	7.186	0.798	0.782	ns	-0.021^{ns}	0.743
ASB17	14	5.009	12.552	0.775	0.805	ns	0.037^{ns}	0.784
ASB2	9	5.122	9.903	0.809	0.809	ns	0.000^{ns}	0.779
ASB23	9	5.220	7.671	0.798	0.813	ns	0.019^{ns}	0.784
HMS3	7	5.024	7.000	0.775	0.805	*	0.038^{ns}	0.772
HMS6	6	3.389	6.312	0.764	0.709	ns	-0.078^{ns}	0.674
HMS7	6	3.431	6.911	0.64	0.713	ns	0.102^{ns}	0.666
HTG10	9	7.120	9.660	0.854	0.864	ns	0.012^{ns}	0.844
HTG4	6	3.366	6.313	0.685	0.707	ns	0.030^{ns}	0.657
HTG6	6	2.115	5.987	0.517	0.530	ns	0.025^{ns}	0.486
HTG7	4	2.192	4.000	0.584	0.547	ns	-0.069^{ns}	0.505
LEX33	8	5.074	8.636	0.798	0.807	ns	0.012^{ns}	0.776
VHL20	9	5.996	9.799	0.809	0.838	ns	0.035^{ns}	0.813
Average	7.714	4.437	7.850	0.745	0.751	ns	0.008^{ns}	-
SE	0.633	0.374	0.57	0.027	0.027	-	-	-
SERT-PE (n=185)								
AHT4	9	5.304	7.757	0.838	0.814	ns	-0.030^{ns}	0.785
AHT5	9	4.520	7.183	0.768	0.781	ns	0.017^{ns}	0.745
ASB17	15	6.076	12.606	0.87	0.838	ns	-0.039^{ns}	0.818
ASB2	12	5.951	9.476	0.773	0.834	ns	0.074^{ns}	0.811
ASB23	8	5.610	8.157	0.822	0.824	ns	0.003^{ns}	0.797
HMS3	7	5.227	6.998	0.762	0.811	ns	0.060^{ns}	0.784
HMS6	7	2.814	6.381	0.67	0.646	*	-0.037^{ns}	0.615
HMS7	8	3.738	6.693	0.676	0.734	ns	0.080^{ns}	0.688
HTG10	10	7.090	9.652	0.843	0.861	ns	0.021^{ns}	0.843
HTG4	7	3.681	6.273	0.778	0.730	ns	-0.066^{ns}	0.687
HTG6	7	2.892	6.064	0.605	0.656	ns	0.077^{ns}	0.598
HTG7	4	2.405	4.274	0.578	0.586	ns	0.013^{ns}	0.538
LEX33	10	5.172	8.369	0.805	0.809	ns	0.004^{ns}	0.780
VHL20	10	6.742	9.856	0.816	0.854	ns	0.044^{ns}	0.835
Average	8.786	4.802	7.839	0.758	0.770	ns	0.016^{ns}	-
SE	0.705	0.399	0.555	0.024	0.023	-	-	-

SP= subpopulation; n= sample size; Na= number of alleles per locus; Ne= effective number of alleles; Ar= allelic richness; Ho and UHe= observed and unbiased expected heterozygosities; HWE= Hardy-Weinberg equilibrium; F_{is}= inbreeding coefficient value; PIC= polymorphism information content; SE= standard error; ns= no significant ($p<0.05$).

significant. Coefficient of gene differentiation value (G_{st}) among the subpopulations was 0.009. According to AMOVA the Φ_{pt} value was 0.010 (p=0.001) and 99 % of variance was explicated within subpopulations and just 1 % among subpopulations.

The value of pairwise F_{st} **(table II)** between AGRE-PE and NCEN-PI of 0.008 (p=0.001) was the highest among the comparisons of the 4 subpopulations of

Table II. *Pairwise subpopulation F_{st} values matrix of the 4 remnant subpopulations Nordestino horse breed. Fst value are below principal diagonal and probability values based on 999 permutations are above principal.* (Matrix dos valores de F_{st} para as 4 subpopulações de remanescentes do cavalo Nordestino. Os valores de F_{st} estão abaixo da diagonal principal e os valores das probabilidades acima da diagonal principal ambos os valores foram obtidos através de 999 permutações).

	AGRE-PE	JUAZ-BA	NCEN-PI	SERT-PE	ARAB
AGRE-PE	-	0.008	0.001	0.035	0.001
JUAZ-BA	0.006	-	0.025	0.006	0.001
NCEN-PI	0.008	0.004	-	0.001	0.001
SERT-PE	0.002	0.004	0.006	-	0.001
ARAB	0.069	0.070	0.072	0.068	-

Nordestino horse, followed by the F_{st} value between AGRE-PE and JUAZ-BA of 0.006 (p=0.008) and NCEN-PI and SERT-PE of 0.006 (p=0.001). Subpopulations from AGRE-PE and SERT-PE had the smallest coefficient of genetic differentiation (F_{st}=0.002; p=0.035). Overall, when all loci were considered, the F_{st} among subpopulations of Nordestino horses was 0.005 (p= 0.001). In other words, the variation among the studied subpopulations corresponded to 0.5 %. The estimated number of migrants between AGRE-PE and JUAZ-BA was around 40, AGRE-PE and NCEN-PI was 31, AGRE-PE and SERT-PE was 110, JUAZ-BA and NCEN-PI was 60, JUAZ-BA and SERT-PE was 56 and NCEN-PI and SERT-PE was 43. There were 3 migrants between ARAB and the subpopulations of Nordestino horses.

Bayesian analysis with MCMC was carried out to study the structure of the four subpopulations of Nordestino horses. There were two groups found according to Bayesian method of the STRUCTURE software. The computing method of Evanno *et al*. (2005) detected that best K value was 2 **(figure 1)**. In the first group were assigned 93.3 % of animals from AGRE-PE, 91.8 % from JUAZ-BA, 95.9 % from NCEN-PI, 95.7 % from SERT-PE and 3.3 % from ARAB.

BOTTLENECK EFFECTS

According to Mode Shift Indicator method, the four subpopulations have not undergone a recent bottleneck event as the L-shaped curves were normal. There was no evidence of bottleneck effects with the Stepwise Mutation Model (SMM) method

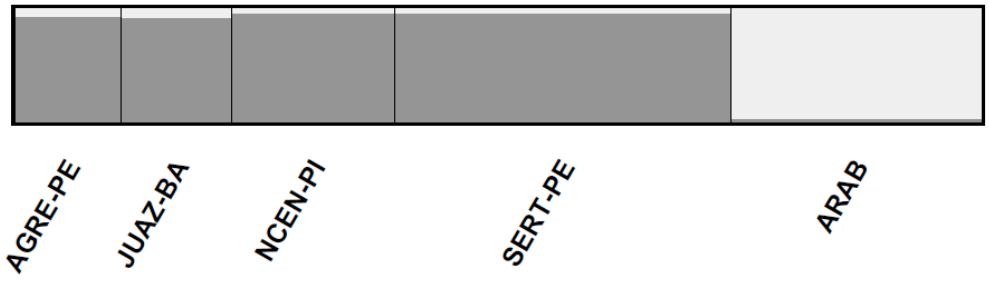

Figure 1. *Structure of remaining from Nordestino horse breed and Arab outgroup for k=2.* (K=2 para a estrutura do remanescente da raça equina Nordestino e a raça Árabe utilizada como outgroup).

or Sign Test (*ST*) and Wilcoxon Rank Test (*WRT*) for all subpopulations of Nordestino horse **(table III)**. However, the TPM (Two-Phase Model) and IAM (Infinite Allele Model) methods evidenced a significant bottleneck effect on the remnant subpopulations of Nordestino horse breed for *ST* and *WRT*.

DISCUSSION

GENETIC DIVERSITY

Genetic diversity corresponds to the variety of alleles presents in a group and can be described by average number of different alleles, allelic richness and expected heterozygosity. Average number of alleles per subpopulation and number of alleles per microsatellite by subpopulation indicated high allelic diversity. All other microsatellite markers tested fit the recommendation of Barker (1994): loci with a minimum of 4 different alleles. The allelic richness is a measure that is not dependent on sample size and among the subpopulations all loci demonstrated values equal to 4 or greater. This result confirmed a high genetic diversity among the subpopulations studied.

UHe was high in the four subpopulations of Nordestino horses. This probably is because the breed is of recent origin and the Nordestino horses have been crossed with other breeds. Established breeds tend to have *UHe* lower than newly formed breeds because the established breed has undergone some inbreeding and selection for the breed specific traits. The highest *UHe* average value was observed for AGRE-PE and SERT-PE. Probably this is a result of crossbreeding, mainly with animals from AGRE-PE with nearby specialized breeds and for SERT-PE due to indiscriminate crossbreeding with horse breeds used in sports practice. Deviations from HWE in isolated loci for remnant subpopulations of Nordestino horses may be due to crossbreeding, negative-assortative mating or gene flow. According to the F_{is} parameter there was no significant deficit of heterozygotes which indicates that the four subpopulations are randomly mating predominantly.

Table III. *Summary of bottleneck test in the 4 remnant subpopulations of Nordestino horse breed.* (Resumo dos resultados obtidos para o teste de efeito de gargalo genético para as 4 subpopulações de remanescentes do cavalo Nordestino).

Subpopulations	IAM	TPM	SMM
AGRE-PE			
ST	Exp.=8.32 (p=0.0007*)	Exp.=8.32 (p=0.0072*)	Exp.=8.25 (p=0.3377^{ns})
WRT	0.00003*	0.00015*	0.78687^{ns}
JUAZ-BA			
ST	8.32 (p=0.0361*)	8.33 (p=0.0369*)	8.23 (p=0.1734^{ns})
WRT	0.00015*	0.04529*	0.87939^{ns}
NCEN-PI			
ST	8.25 (p=0.0006*)	8.27 (p=0.0343*)	8.30 (p=0.0653^{ns})
WRT	0.00003*	0.01477*	0.57239^{ns}
SERT-PE			
ST	8.22 (p=0.0006*)	8.29 (p=0.0069*)	8.26 (p=0.16874^{ns})
WRT	0.00003*	0.00015*	0.29578^{ns}

IAM= infinite allele model; TPM= two-phase model; SMM= stepwise mutation model; ST= sign test (number of loci with heterozygosity excess); WRT= wilcoxon rank test with probability of heterozygosity excess; Exp.= expected number of loci with heterozygosity excess; *p<0.05; ns= no significant;

GENETIC DIFFERENTIATION

G_{st}, F_{st} and Φ_{pt} values among the four subpopulations of Nordestino horses showed that the animals were a single genetic group with no evidence of genetic differentiation among them. Giacomoni *et al*. (2008) analyzed three subpopulations (Ipiranga, Nova Esperança and Promissão) of Pantaneiro horses and Cothran *et al*. (2011) studied three subpopulations (Apure, Aragua and Mérida) of Venezuelan Criollo and also observed low genetic differentiation, no evidence of genetic differentiation among horses of the same breed from different places. That the subpopulations of Nordestino horse were the same gene pool was corroborated by the distribution of the animals in the STRUCTURE analysis where only two defined groups (k=2): one the Arab outgroup and other the Nordestino horses. Costa *et al*. (1974) reported Bahia, Pernambuco, Ceará and Piauí states were the places where the Nordestino horse was essentiality formed and there was a larger herd at that time. There has been a fair number of horses selected from those states to integrate into the foundation of the stud nucleus. Until the last years of the ABCCN animals were widely transported among states and farms both intentionally and due to natural wandering. In the 1990s the Nordestino horse breed Association stopped animal registrations, but even so the Nordestino horse survives in the Caatinga biome from the Brazilian Northeast region.

The horses have exchanged genes among the subpopulations with a large number of individual migrants between AGRE-PE and SERT-PE (110 migrants estimated) followed by NCEN-PI and JUAZ-BA (60 migrants estimated) which has served to homogenize the population. The high gene flow estimated among subpopulations and the historical accounts demonstrate that the populations represent a single interbreeding genetic group.

BOTTLENECK EFFECT

There was no strong evidence of a recent population bottleneck which was consistent with the high genetic variability within each subpopulation. Probably, the bottleneck was not recent, or just a demographic bottlenecks occurred without genetic one or/and the subpopulations were not completely isolated and contains genes from migrants that had disguised genetic effects of the bottlenecks (Luikart *et al*., 1998a,b). The Nordestino horse breed may be experienced a gradual decline in their overall population numbers mainly because of closure of the Association due to lack of interest in the breed, the current castration practice on the males, the mechanization process and increase of automobile in rural areas. Those results and reasons may be suggested that a demographic bottleneck event have occurred in the Nordestino horse breed. However it was not enough to cause genetic differentiation how was detect and discussed in the topic above.

REFERENCES

ABCCN. 1987. Associação Brasileira dos Criadores do Cavalo Nordestino. Regulamento do registro genealógico do cavalo Nordestino. ABCCN. Recife. pp. 33-34.

Barker, J.S.F. 1994. A global protocol for determining genetic distances among domestic livestock breeds. International Committee for World Congresses on Genetics Applied to Livestock Production. V World Congress on Genetics Applied to Livestock Production. Guelph. Canada. pp. 501-508.

Belkhir, K.; Borsa, P.; Chikhi, L.; Raufaste, N. and Bonhomme, F. 2003. Laboratoire genome, populations, interactions: CNRS UMR. Université de Montpellier II. Montpellier. France.

Coelho, E.G.A.; Oliveira, D.A.A.; Teixeira, C.S.; Sampaio, I.B.M.; Rodrigues, S.G. e Alves, C. 2004. Comparação entre métodos de estocagem

de DNA extraído de amostras de sangue, sêmen e pêlos e entre técnicas de extração. *Arq Bras Med Vet Zoo*, 56: 111-115.

Cornuet, J.M. and Luikart, G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144: 2001-2014.

Costa, N.; Val, L.L. e Leite, G.U. 1974. Estudo da preservação do cavalo Nordestino. Departamento de Produção Animal. V. 38. Recife. Brasil. pp. 1-15.

Cothran, E.G.; Canelon, J.L.; Luis, C.; Conant, E. and Juras, R. 2011. Genetic analysis of the Venezuelan Criollo horse. *Genet Mol Res*, 10: 2394-2403.

Evanno, G.; Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*, 14: 2611-2620.

Falush, D.; Stephens, M. and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164: 1567-1587.

Falush D.; Stephens, M. and Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes,* 7: 574-578.

FAO/ISAG. 2011. Food and Agriculture Organization of the United Nations and International Society for Animal Genetics. Molecular genetic characterization of Animal Genetic Resources. Guidelines. FAO. Rome. 100 pp.

FAO. 2012. Food and Agriculture organization of the United Nations. Draft guidelines on *in vivo* conservation of Animal Genetic Resources. FAO. Rome. 160 pp.

Giacomoni, E.H.; Fernández-Stolz, G.P. and Freitas, T.R.O. 2008. Genetic diversity in the Pantaneiro horse breed assessed using microsatellite DNA markers. *Genet Mol Res*, 7: 261-270.

Goudet, J. 1995. FSTAT (vers. 1.2): a computer program to calculate F-statistics. *J Hered*, 86: 485-486.

Hoffmann, I.; Ajmone Marsan, P.; Barker, J.S.F.; Cothran, E.G.; Hanotte, O.; Lenstra, J.A.; Milan, D.; Weigend, S. and Simianer, H. 2004. New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a joint ISAG/ FAO working group. In: Proceedings of 29th International Conference on Animal Genetics. Tokyo. Japan. 123 pp.

Jakobsson, M. and Rosenberg, N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23: 1801-1806.

Kalinowski, S.T.; Taper, M.L. and Marshall, T.C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*, 16: 1099-1106.

Lang, R. 2006 GhostView version 4.9 software.

Luikart, G.; Allendorf, F.W.; Cornuet, J.M. and Sherwin, W.B. 1998a. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J Hered*, 89: 238-247.

Luikart, G.; Sherwin, W.B.; Steele, B.M. and Allendorf, F.W. 1998b. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Mol Ecol*, 7: 963-974.

Melo, J.B. 2011. Caracterização zoométrica do remanescente da raça equina Nordestina nos estados de Pernambuco e Piauí. Doutorado em Zootecnia thesis. Universidade Federal Rural de Pernambuco. Recife.

Ott, J. 1992. Strategies for characterizing highly polymorphic markers in human gene mapping. *Am J Hum Genet*, 51: 283-290.

Peakall, R. and Smouse, P.E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes*, 6: 288-295.

Pritchard, J.K.; Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959.

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes*, 4: 137-138.

Rousset, F. 2008. Genepop'007: a complete reimplementation of the genepop software for Windows and Linux. *Mol Ecol Resour*, 8: 103-106.

Weir, B.S. and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370.

Wright, S. 1951. The genetical structure of populations. *Ann Eugenics*, 15: 323-354.