African Journal of Biotechnology Vol. 9 (3), pp. 293-299, 18 January, 2010
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2010 Academic Journals

Full Length Research Paper

Analysis of genetic diversity and estimation of inbreeding coefficient within Caspian horse population using microsatellite markers

Hosain Shasavarani and Ghodrat Rahimi-Mianji*

Laboratory for Molecular Genetics and Animal Biotechnology, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, P.O. Box, 578 Sari, Iran.

Accepted 21 September, 2009

The present study was undertaken to genetically evaluate Caspian horses for genetic diversity and to asses whether they have experienced recent population bottlenecks. A total of 100 individuals were characterized for within breed diversity using 16 microsatellite markers. The estimated mean number of alleles was 8.69 per locus, with a total of 139 alleles in the genotyped sample. The mean effective number of alleles in the Caspian horse population was 5.86, ranging from 3.49 to 8.49. The average observed heterozygosity in the present study (0.52) was lower than to the expected heterozygosity (0.82), which may reflect the narrow genetic base of the current population of this breed. All marker loci employed in this study were very informative with an average of 0.80. The Chi-square and likelihood ratio tests performed to examine population for HWE showed some highly significant deviations from HWE. Estimated values of Wright's fixation index, F_{IS} (0.367) indicates a certain level of heterozygote deficiency. A significant heterozygote excess on the basis of different models, as revealed from Sign and Wilcoxon rank test suggested that Caspian horse population is not in mutation-drift equilibrium. But, the Mode-shift indicator test showed a normal 'L' shaped distribution for allelic class and proportion of alleles, thus indicating the absence of bottleneck events in the recent history of this breed. The present work is a contribution to the knowledge of population structure and to the assessment of genetic diversity that may be helpful to horse breeders in designing and managing breeding or conservation strategies for the Caspian horse breed.

Key words: Genetic diversity, microsatellite markers, Caspian horse breed.

INTRODUCTION

The Caspian is thought to be one of the oldest horse breeds in the world, going back to 3000 BCE. These horses are quite small, typically ranging from 9 - 10 hand height. This characteristic stands them below the standard height for a horse, but they are not considered ponies because of their physical appearance and history. The original Caspian horse distribution spans between the southern coast of Caspian Sea (Gilan, Mazandaran, Golestan provinces) and northern areas of Alburz mountain, although in recent years a number of individuals are raised in other parts of the country. These perfect little horses are very similar to the Arabian horses. The Caspian horse is also intelligent, highly alert and very friendly, making well trained horses suitable for younger riders in show jumping and racing (Figure 1). Ancient stone engravings and archeological findings point to the fact that the Caspian horse may in truth be the first horse to be used in the Middle East. This horse has had much success abroad and is now bred in Bermuda, England and New Zealand. The current estimated registered population of Caspian horses in Iran is approximately 150 individuals (Seyedabadi et al., 2006). Small populations are exposed to loss of genetic variability which can lead to reduced fitness, inbreeding depression and in some cases, extinction of the breed.

The development of new genetic tools has brought about great advances in individual recognition and DNA

*Corresponding author. E-mail: rahimimianji@yahoo.com. Tel: +98 151 3822565. Fax: +98 151 3822577

Figure 1. A typical Caspian horse (copyright MobarakAndish Institute).

markers such as microsatellites have proved to be useful in clarifying population structure and evaluating genetic diversity (Marletta et al., 2006; Plante et al., 2007; Avdi and Banos, 2008). The present study was undertaken to genetically evaluate Caspian horses for genetic variability and to asses whether they have experienced any recent genetic bottlenecks based on the analysis of micro-satellite markers.

MATERIALS AND METHODS

Blood samples and genomic DNA isolation

Blood samples were collected from 100 animals at five different breeding locations of Caspian horse breed as shown in Figure 2. About 5 ml of blood per animal was collected in EDTA (0.5 mM, pH. 8.0) coated vacationers. Genomic DNA was isolated from blood using the high salt extraction procedure (Montgomery and Sise, 1990).

Microsatellite loci and PCR based profiling

A set of sixteen microsatellite markers (VHL20, HTG4, HTG7, HTG10, AHT5, ASB2, HMS2, HMS6, HMS7, NVHEQ43, TKY335, TKY343, UMNe61, UM026, UM040, UM042) were genotyped in this study. Nine of these markers were those recommended for parentage testing by the International Society for Animal genetics (ISAG) Equine Genetics Standing Committee. The primer sequences and the annealing temperatures for each primer are presented in Table 1. Genomic DNA was amplified by PCR and each 25 µl reaction tube consisted of DNA (about 50 ng), primers (50 pmol each), dNTP (200 µM each), 10× buffer (10 mM Tris, 50 mM KCl, 0.1% gelatin, pH. 8.4), $MgCl_2$ (1.5 mM) and *Taq* DNA polymerase (1 U). The PCR products were resolved by electrophoresis through 8% denaturing urea-polyacrylamide gel and visualized by silver staining. The alleles were scored manually from the silver staining gel.

Computation and statistical analysis

Allelic frequency was estimated using genotype counting. The

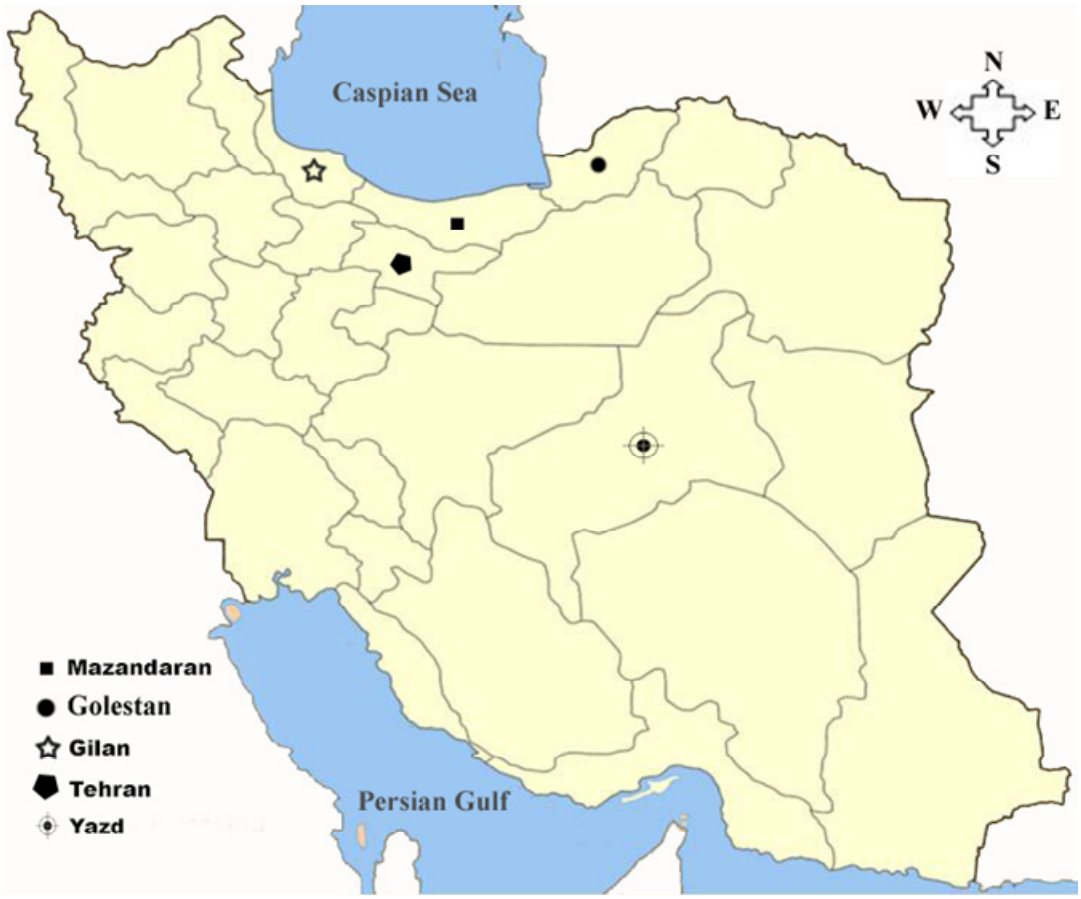

Figure 2. Geographical location of the Caspian horse population sampled in the present study

within breed genetic variation parameters observed and expected heterozygosity, observed and effective number of alleles at each microsatellite marker locus were calculated using the POPGENE computer program version 1.31 (Yeh et al., 1999). Based on allelic frequencies, expected heterozygosity levels were calculated for each microsatellite as described by Nei (1973). The effective number of alleles was calculated according Kimura and Crow (1964). The probability of random mating in the population was estimated by both Chi-square (χ^2) and likelihood ratio tests (G^2) to examine Hardy-Weinberg equilibrium (HWE) at each locus. Inbreeding coefficient *(F_{IS})* per locus and for the population (heterozygote deficiency) was calculated as the difference between observed and expected heterozygosity (Nei and Kumar, 2000). Polymorphic information content *(PIC)* value of the microsatellite loci was calculated on the basis of observed allele frequencies (Botstein et al., 1980), using Cervus software (Marshall, 2000). Bottleneck events in the population were tested by two methods. The first method consisted in the application of three excess heterozygosity tests developed by Cornuet and Luikart (1996). The probability distribution was calculated using 1000 simulations under three models of evolution: infinite allele model (IAM), stepwise mutation model (SMM) and two phase model of mutation (TPM). The second method is the graphical representation of the model shift indicator proposed by Luikart et al. (1998). Loss of rare alleles in bottlenecked populations is detected when one or more of the common allele classes have a higher number of alleles than the rare allele class. These two methods were conducted using Bottleneck computer program (Piry et al., 1999).

RESULTS AND DISCUSSION

In the present study, all marker loci were observed to be polymorphic and a total of 139 distinct alleles were detected across 16 microsatellites in Caspian horse population. PCR product size range varied from 85 - 107 at marker locus VHL20 to 248 - 270 at locus UM040. The number of alleles ranged from 5 (UM042) to 12 (VHL20), with a mean of 8.69 per locus (Table 2). This mean number of alleles per microsatellite locus in Caspian horses was higher than that reported in other Asian (6.58 - 7.75), European (4.50 - 7.08), South American (5.67-7.67) and North American (6 -7.25), horse breeds (Luis et al., 2007) and lower than (9.58) that reported in the Italian Murgese horse breed (Pieragostini et al., 2005). The differences in average number of observed alleles may be attributed to different set of microsatellite markers, number of markers, population structure as well as different horse breeds. In our study, the observed heterozygosity values across the 16 polymorphic marker loci ranged from 0.32 (TKY335) to 0.95 (NVHEQ43), with a mean of 0.52. The expected heterozygosity varied from 0.71 (HTG4) to 0.88 (UMNe61), with a mean of 0.82. The mean observed and expected heterozygosity values across 16 microsatellite

Table 1. Primer sequences, chromosome numbers and annealing temperatures for 16 microsatellie loci employed for genetic characterization of Caspian horse.

Locus	Primer sequences (5′--------3′)	Chromosome location	Annealing temperature (°C)
AHT5	F-ACGGACACATCCCTGCCTGC R-GCAGGCTAAGGGGGCTCAGC	8	59
ASB2	F-CCACTAAGTGTCGTTTCAGAAGG R-CACAACTGAGTTCTCTGATAGG	15	60
HMS2	F-CTTGCAGTCGAATGTGTATTAAATG R-ACGGTGGCAACTGCCAAGGAAG	10	59
HMS6	F-GAAGCTGCCAGTATTCAACCATTG R-CTCCATC TTGTGAAGTGTAACTCA	4	58
HMS7	F-CAGGAAACTCATGTTGATACATC R-TGTTGTTGAAACATACCTTGACTGT	1	60
HTG4	F-CTATC TCAGTCTTGATTGCAGGAC R-CTCCCTCCCTCCCTCTGTTCTC	9	60
HTG7	F-CCTGAAGCAGAACATCCCTCCTTG R-ATAAAGTGTCTGGGCAGAGGCTGCT	4	60
HTG10	F-CAATTCCCGCCCCACCCCCGGCA R-TTTTTATTCTGATCTGTCACATTT	21	59
TKY335	F-TGAGAGTAGTAACCTGCATC R-TAGGCAAGTCTCAGTTTTCC	2	55
TKY343	F-TAGTCCCTATTTCTCCTGAC R-AAACCCACAGATACTCTAGA	11	55
NVHEQ43	F-TGACACAAGATAAAAGCCCCAGG R-GATTGGGAAAAGAGCACAGCC	25	60
UMNe61	F-CCGCCATCATTACTAAGCAG R-TCTTTTCCTCCAATCCTACTCC	-	60
UM026	F-CCCAAAATCAATTAGGTCTC R-ATCAGTTGCTCTCTACTTTT	1	58
UM040	F-CTCTTGTACATGTCTCCTTG R-TACTTTCTCTCTTCCAAACC	10	56
UM042	F-GGCATCCCACATACAAAG R-GAAGCAACAGTCAATTCAG	16	56
VHL20	F-CAAGTCCTCTTACTTGAAGACTAG R-AACTCAGGGAGAATCTTCCTCAG	30	60

loci in Sorraia horse breed ranged from 0.45 to 0.46, respectively (Luis et al., 2007) and 0.63 in Greek Skyros breed (Avdi and Banos, 2008). The mean effective number of alleles in the Caspian horse population was 5.86, ranging from 3.49 (HTG4) to 7.60 (ASB2). The average observed heterozygosity at the present study (0.52) was lower than to the expected heterozygosity calculated (0.82) from allelic diversity, which may reflect the narrow genetic base of the current population of this breed.

Genetic markers showing PIC values higher than 0.5 are normally considered as informative in population genetic analyses (Botstein et al., 1980). All marker loci employed in this study were informative since all PIC values were well above this level (Table 2). The average observed allele number and PIC values obtained in our study were higher than previous results reported for Caspian horse breed (Seyedabadi et al., 2006; Amirnia et al., 2007). The obtained results in both previously reports has been analyzed on the basis of the same number of samples and also with the same microsatellite markers (except one). These differences could probably be explained by the higher number and/or the choice of the markers, higher number of samples and also wide ranged sampling localities in the present study. The obtained results in our study may be more reliable than the two previous reports because of the addition of further loci, locations, herds as well as additional animals in resolving the closer relationships between animals in Caspian horse population. Chi-square and likelihood ratio tests performed to examine population for HWE showed highly significant deviations from HWE ($p < 0.001$). Only two out of sixteen microsatellite loci, HTG10 and NVHEQ43,

Table 2. Allele size, number of alleles (observed N_o, effective N_e, observed heterozygosity *(H_o)*, expected heterozygosity *(H_e)*, polymorphism information *(PIC)* and Wright's fixation index estimates *(F_{IS})* at 16 microsatellite markers in Caspian horse.

Locus	Allele size range (bp)	N_o	N_e	H_o	H_e	PIC	F_{IS}
AHT5	118 - 142	11	6.43	0.73	0.84	0.83	0.141
ASB2	154 - 188	11	7.60	0.51	0.87	0.85	0.402
HMS2	214 - 238	11	5.79	0.52	0.83	0.81	0.351
HMS6	157 - 173	9	5.97	0.45	0.83	0.81	0.443
HMS7	170 - 186	9	4.39	0.41	0.77	0.74	0.451
HTG4	129 - 141	6	3.49	0.42	0.71	0.68	0.412
HTG7	116 - 130	8	6.91	0.49	0.86	0.84	0.402
HTG10	98 - 114	8	4.38	0.81	0.77	0.74	-0.061
TKY335	249 - 263	7	6.42	0.32	0.84	0.82	0.621
TKY343	149 - 171	8	4.49	0.43	0.78	0.76	0.434
NVHQ43	133 - 161	11	6.92	0.95	0.86	0.85	-0.112
UMNE61	121 - 145	9	8.49	0.72	0.88	0.87	0.164
UMO26	204 - 220	7	5.23	0.33	0.81	0.78	0.562
UMO40	248 - 270	7	5.48	0.59	0.82	0.79	0.242
UMO42	216 - 226	5	4.68	0.14	0.79	0.75	0.853
VHL20	85 - 107	12	7.13	0.42	0.86	0.84	0.562
Mean ± Sd	-	8.69 ± 2.1	5.86 ± 1.4	0.52 ± 0.20	0.82 ± 0.05	0.80 ± 0.05	0.367 ± 0.25

showed higher observed heterozygosity than the expected values, whereas for the rest of the loci observed heterozygosity was lower than expected (Table 2). The condition of strong HW disequilibrium found for the majority of the loci analyzed (P value is statistically significant for 14 out of 12), is probably attributed to the small population size, selection at or near the genomic locus, non-random mating, inbreeding, or genetic drift. In addition, the possible occurrence of null alleles could have lead to false observation of homozygotes which could account for more deviations from HWE. A broad variation was found for the F_{IS} index value among loci, ranging from -0.112 (NVHQ43 locus) to 0.853 (UM042 locus). The overall F_{IS} value, among all loci was significantly higher than zero (0.367) indicating a certain level of heterozygote deficiency. All the 16 examined microsatellite markers, except HTG10 and NVHEQ43 contributed to this observed heterozygote shortage (Table 2). This may be due to the fact that small number of individuals involved in reproduction or in the last few decades mating has been occurred among closely related animals. The heterozygote deficiency could also be explained as a Wahlund effect if population subdivision is occurring, linkage with loci under selection (genetic hitchhiking), population heterogeneity, null alleles (non-amplifying alleles) or inbreeding. Positive F_{IS} value suggested inbreeding to be one of the main causes for shortage of heterozygotes in Caspian horse population.

In the last years, Caspian horse has become increasingly popular in other parts of the country as well and the geographical distances may lead to subdivision of the breed. To characterize the demographic bottlenecks the Sign and Wilcoxon sign rank tests were utilized. In these two tests, under three models of microsatellite evolution, the expected number of loci with heterozygosity excess were 16, 12 and 16 for IAM, TPM and SSM models, respectively. The probability value obtained from Sign test was 0.00023 (IAM), 0.00022 (TPM) and 0.1446 (SSM) and thus rejects the null hypothesis for both IAM and TPM models, indicating that the population has undergone a genetic bottleneck. Using the Wilcoxon rank test the probability values were 0.00002 (IAM), 0.00002 (TPM) and 0.033 (SMM), indicating that the null hypothesis also should be rejected in all three models, thus also demonstrating that the Caspian horse population has experienced a genetic bottleneck. The Mode-shift indicator test was also utilized as a second method to detect potential bottlenecks, as the non-bottlenecked populations that are near mutation-drift equilibrium are expected to have a large proportion of alleles at low frequencies. In the Mode-shift indicator test the microsatellite alleles are organized into 10 frequency classes, which permit checking whether the distribution followed the normal L-shaped form, where alleles with low frequencies are the most numerous. The observed allele frequency distribution at the present study (Figure 3) indicates the absence of bottleneck events in the recent history of Caspian horse population. When analyzing allele frequency distributions to test for bottlenecks, it may be necessary to assume that the test population is random mating, has no substructure, has no recent immigration, loci are neutral, and that sampling is representative of the population (Luikart et al., 1998). In the present study, both χ^2 and G^2 tests showed highly signifycant deviations from HWE ($p<0.001$) in Caspian horse population. Therefore it can be stated that the observed

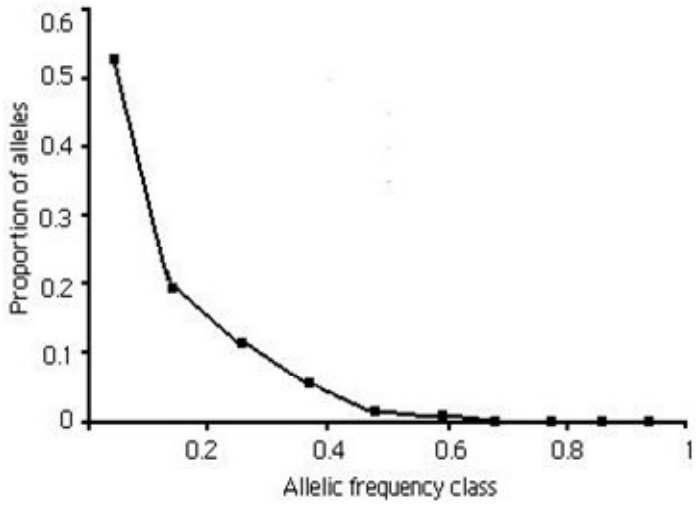

Figure 3. Normal L-shaped curve of distribution of proportion of alleles in different allelic frequency classes

bottleneck status in Caspian horse population has been occurred in the past history of this breed. This hypothesis can be confirmed by the Mode-shift indicator test, thus indicating the Caspian horse population has not experienced a recent genetic bottleneck. It has been reported Mode-shift indicator model with and/or without loci deviating from Hardy–Weinberg proportions in the datasets from which genotype frequency data was available, made no difference in the test results for bottleneck detection (Luikart et al., 1998). Moreover, it has been shown that the bottlenecked populations may not have a mode-shifted allele distribution for the following reasons: (1) the bottleneck was not recent or small enough to be detectable, (2) not enough polymorphic loci and/or individuals were sampled to have sufficient power for detecting the bottleneck, (3) the individuals sampled were not representative of the bottlenecked population, (4) a demographic bottleneck occurred but not a genetic bottleneck, and (5) the bottlenecked population is not completely isolated and contains genes from immigrants that have obscured the genetic effects of the bottleneck (Luikart et al., 1998). Therefore, if analyses of allele frequency distributions fail to detect a mode-shifted distribution of allele frequencies, one should not conclude that a population has not been bottlenecked; one can only conclude that a bottleneck is not likely to have occurred in the recent past (Luikart et al., 1998). According to the analyzed data in the present study, the population size of Caspian horse may has decreased to low numbers in the past, but has recently increased due to breeding in different parts of the country.

Conclusion

The present work contributes to the knowledge of population structure and assessment of existing genetic diversity in the Caspian horse population. Considering the results from all three test models from Sign and Wilcoxon rank test, it is clear that serious demographic bottlenecks have most probably occurred in this breed. But, the Mode-shift indicator test showed a normal 'L' shaped distribution for allelic class and proportion of alleles, thus indicating the Caspian horse population has not experienced a recent bottlenecks and has been occurred in the past history of this breed. In the future, genetic analysis for other Iranian horse breeds and their comparisons need to be carried out to determine the phylogenic evolutionary relationships and genetic distances among the indigenous Iranian equine breeds. However, for Caspian horse breed, high priority action is necessary considering that the breeding practices exercised by local horse owners, may further weaken the diversity levels through the breeding of relatives. To make a start, providing pedigrees for Caspian horse breeders to collect correct information for breeding management is an obvious strategy. When both genealogical and molecular information is available, it can be combined to calculate the coancestry conditional on markers which might better help to estimate the genetic structure of Caspian horse population.

ACKNOWLEDGMENTS

This work would not have been possible without the collaboration and support of horse owners. Authors are particularly grateful to all horse owners for their helpful, cooperation and kind assistance during sample collection. Funding for this study was provided by the research department of Sari Agricultural Science and Natural Resources University.

REFERENCES

Amirnia C, Seyedabadi M, Banabazi MH, Kamali MA (2007). Bottleneck study and genetic structure of Iranian Caspian horse population using microsatellites. Pakistan J. Biol. Sci., 10 (9): 1540-1543.

Avdi M, Banos G (2008). Genetic diversity and inbreeding in the Greek Skyros horse. Livestock Science, 114: 362-365.

Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am. J. Human Genet. 32: 314-331.

Cornuet JM, Luikart G (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001–2014.

Kimura M, Crow JF (1964). The number of alleles that can be maintained in a finite population. Genetics. 61: 763-771.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. J. Hered. 89: 238–247.

Luis C, Cothran EG, Oom MM (2007). Inbreeding and genetic structure in the endangered Sorraia horse breed: Implications for its conservation and management. J. Heridity 98 (3): 232-237.

Marletta D, Tupac-Yupanqui I, Bordonaro S, Garcia D, Guastella AM, Crisione A, Canon J, Dunner S (2006). Analysis of genetic diversity and determination of relationships among western Mediterranean horse breeds using microsatellite markers. J. Anim. Breed. Genet.

123:315–325.
Marshall TC (2000). A program designed for a large-scale parentage analysis using codominance loci (CERVUS, version 2.0). University of Edinburgh, UK.
Montgomery GW, Sise JA (1990). Extraction of DNA from sheep white blood cells. New Zealand J. Agric. Res. pp. 337-441.
Nei M (1973). Analysis of gene diversity in subdivided populations. Pro. Natl. Acad. Sci. U.S.A. 70: 3321-3323
Nei M, Kumar S (2000). Molecular evolution and phylogenetics. Oxford Uinversity press, pp. 236
Pieragostini E, Rizzi R, Bramante G, Rosati A, Perrotta G, Caroli A (2005). Genetic study of Murgese horse from genealogical data and microsatellites. Ital. J. Anim. Sci., 4: 197-202.
Piry S, Luikart G, Cornuet JM, (1999). Bottleneck: a computer program for detecting recent reductions in effective population using allele frequency data. J. Hered. 90: 502–503.
Plante Y, Luis J, Lucas Z, Coling D, March BD, Buchanan F (2007). Genetic diversity in a Feral horse population from Sable Island, Canada. J. of Heredity, 98 (6): 594–602.
Seyedabadi H, Amirnia C, Banabazi MH, Emrani H (2006). Parentage verification of Iranian Caspian horse using microsatellite markers. Iranian J. Biotechnol., 4 (4): 260-264.
Yeh FC, Yang RC, Boyle T (1999). POPGENE, version 1.31. Microsoft windows-based Freeware for population genetic analysis. Molecular Biology and Technology Center, University of Alberta, Canada