

## GENETIC DIVERSITY, POPULATION STRUCTURE AND PHYLOGENETIC RELATIONSHIP OF RACE, SPORTS, DRAFT AND WILD TYPE HORSES

Iqra Atiq<sup>1</sup>, Imran Zahoor<sup>1</sup>, A. Basheer<sup>1,\*</sup> and Waseem Khan<sup>2</sup>

<sup>1</sup>Genetics and Genomics Laboratory, Department of Livestock Production, University of Veterinary and Animal Sciences, Lahore, Pakistan; <sup>2</sup>Department of Wildlife and Ecology, University of Veterinary and Animal Sciences, Lahore, Pakistan

\*Corresponding author's e-mail: atia.basheer@uvas.edu.pk

In the current study, data of 930 horses was used and divided into four types of categories i.e. Draft, Wild, Sports and Race types. A set of ten microsatellite markers HTG10, ASB2, ASB17, HMS3, AHT4, HMS7, VHL20, AHT05, HTG4, and HMS6 were studied. Allelic frequency, total number of alleles, observed number of alleles, expected number of alleles, heterozygosity, Hardy-Weinberg equilibrium was calculated using GeneAIEX 6.502 software. Genetic structure of horse breeds was studied using STRUCTURE & STRUCTURE harvest software, while results were visualized by CLUMP and DISTRUCT software. Moreover, Phylogenetic analysis was performed using PHYLIP and data was visualized in MEGA 7. Maximum number of alleles was observed for sports category, and most polymorphic markers among all types of horses was ASB17. The maximum and minimum values of inbreeding coefficient was found in wild and race horses respectively. Maximum observed heterozygosity was found in case of race horses; however, highest value of expected heterozygosity was found for sports horses. Present study showed that even with limited number of markers considerable level of genetic diversity between and among all types of horses was observed.

**Keywords:** *Equus ferus caballus*, grazing animals, domestic horse, wild populations, inter breeding, genetic markers

### INTRODUCTION

The horse (*Equus ferus caballus*) is one of subspecies of *Equus ferus* and belongs to family *Equidae* which includes Asses, Zebras and Horses. All species are grazing animals that are well known for their strength and speed. Over the past 45-55 million years, horse evolved from *Eohippus*. Other subspecies of *Equus ferus* like tar pan (*Equus ferus ferus*) were undomesticated which are extinct now (Kavar and Dove, 2008). Humans started to domesticate horse about 4000BC ago (Goodwin, 2007) in Eurasian steppes (Petersen *et al.*, 2013). Two alternative hypotheses can be formulated for the origin of the domestic horse from wild populations. A restricted origin hypothesis hypothesizes that the domestic horse was developed through selective breeding from a narrow wild stock and, thereafter, domestic horses were distributed to other regions. According to other hypothesis domestication is a complex process that involves multi-generation selection on traits that permit stable existence of a population. According to this, horses may have been independently captured from diverse wild populations and then increasingly bred in captivity (Vila, 2001). Mitochondrial DNA (mtDNA) sequence of pre-domestic and domestic horses has revealed that European wild horse populations also contributed to the gene pool of domestic horses (Cothran *et al.*, 2011).

The horse had been mainly used for sport, food, commerce, warfare, transportation, religion, pleasure, trade, agriculture, communication, industry, gift, symbol and recreation. Moreover, horses have played an important role in many civilizations and cultures (Bjornstad *et al.*, 2003). Nowadays horses are kept as a hobby and they are considered as companion animal in many sports and recreation. This change has few major reasons: firstly, the population of riding and racing breeds has principally increased, whereas numerous draught breeds are currently in danger of extinction; secondly, horse breeders have powerfully customized their breeding objectives. Moreover, outcrossing is commonly practiced, to improve performance among local and international breeds (Diaz *et al.*, 2002).

Domestic horse genome encompasses 31 pairs of autosomes and one pair of sex chromosome. Whereas, wild horse (Prezwalski) possesses an extra pair of chromosome, yet it can interbreed with domestic horse and produce fertile offspring (Schubert *et al.*, 2014) this may be because of fusion of chromosome 23 and 24 of Prezwalski horse in domestic horse. The horse genome project mapped 2.7 billion base pairs. It was first sequenced in 2006, whereas full map was established in 2009. Horse genome is larger than dog genome but smaller than human and bovine genome (Horse genome project 2009). The first fully sequenced horse genome was thoroughbred mare named "Twilight", in 2006-2007.

Among genetic markers, microsatellite markers are highly

polymorphic and have high mutability. Hence, provide valuable information for constructing linkage maps and evaluating genetic diversity (Miah *et al.*, 2015). Microsatellite markers are also known as short tandem repeats (STRs) or simple sequence repeats (SSRs). Microsatellites are repeats of two to six nucleotides. These are abundant sequences that are present throughout eukaryotic genomic DNA (within intron and exons). They show variability in number of repeats and serve as very useful tool for genetic mapping (Cheng and Crittenden 1994; Giacomoni *et al.*, 2008; Boruszewska *et al.*, 2009; Bech *et al.*, 2010). They are highly polymorphic, stable, and can be co-dominantly inherited (Berber *et al.*, 2014). Moreover, autosomal microsatellites markers can be used for parentage testing, estimation of diversity, calculation of genetic distances, inbreeding coefficient and population admixture (Hanotte and Jianlin, 2005). It can be successfully used to construct phylogenetic tree among different species as well as at among different taxonomic levels (Achmann *et al.*, 2004; Bigi and Perrotta, 2012).

Knowledge about population history, structure, and genetic diversity is very valuable for conservation and management programs of a species, as it provide basic information about level of genetic variation, historical pattern of population subdivision and gene flow (Gandini and Oldenbroek, 1999; Canon *et al.*, 2001; Hanotte and Jianlin, 2005; Vicente *et al.*, 2005; Sarkar *et al.*, 2006). This study aimed at investigating the genetic diversity, population structure and phylogenetic relationship among race, sports, draft and wild type horses.

## MATERIALS AND METHODS

**Ethics statement:** Ethical permission (number DR/151) for the collection of blood samples was obtained from the Ethical review committee for animal research of University of Veterinary and Animal Sciences, Lahore, Pakistan. Current study is based on the data of 930 animals. Blood samples from (n=40) thoroughbred horses were collected from different stud farms of Punjab registered with Jockey Club of Pakistan. Genomic DNA was extracted using QIAamp DNA Blood Kit (QIAGEN, Germany). In addition to 40 samples obtained locally, the genotypic data for the remaining 890 horses were taken from previous studies (Warmuth *et al.*, 2011; Khanshour *et al.*, 2014; Uzans *et al.*, 2015). In the present study, the total number of horses were divided into four types Draft (n=202), Wild (n=253), Sports (n=256) and Race (n=219). Draft type horses includes Belgian, Breton, Clydesdale, Shire, Percheron and Suffolk. Sport type include Andalusian, Morgan, American Saddlebred, Standardbred Trotter, Standardbred Pacer and Paso Fino. Racing type include Akhal Teke, Syrian and Thoroughbred while the wild type horses include Sable Island and Przewalski (Table 1).

**Table 1. Summary break-up of wild, race, draft, and sports horses.**

Horse Types	No. of animals	Temperament	Breeds
Wild	253	Wild	a. Sable Island b. Przewalski
Race	219	Hot bloods	a. Akhal Teke b. Syrian
Draft	202	Cold bloods	c. Thoroughbred (40+50) a. Belgian b. Breton c. Clydesdale d. Shire e. Percheron f. Suffolk
Sports	256	Warm bloods	a. Andalusian b. Morgan horse c. American Saddlebred d. Standardbred Trotter e. Standardbred Pacer f. Paso Fino

Ten microsatellite markers AHT4, AHT5, ASB2, ASB17, HMS6, HMS7, HMS3, HTG4, VHL20, and HTG10 were used to study the genetic diversity and population structure of different types of horses. Information about the loci location, primer sequences and allele sizes can be found in Table 2.

For all markers together, the Fisher's method implemented in GENEALX 6.5 was used to detect the significant deviation of a population from HWE. We calculated genetic diversity indices among race, sports, draft and wild type horses by using GENEALX 6.5. These included the average number of alleles (Na), the effective number of alleles (Ae), Shannon's information index (I), Observed heterozygosity (Ho) expected heterozygosity (He), unbiased expected heterozygosity (uHe), and fixation index (F). The inbreeding coefficient (FIS) was also calculated using 10 000 bootstraps as implemented in GENETIX 4.05 (Belkhir *et al.*, 1996-2004). The genetic relationship of the race, sport, draft and wild types horses was studied by using two methods. First we computed Nei genetic distances and then boot strapping 1000 times across all loci using PHYLIP 3.9 software (Felsenstein, 1989-2006.). After ward neighborhood joining method of PHYLIP was used to construct phylogenetic tree and was visualized by using MEGA7 (Tamura *et al.*, 2013). In the second method, the genetic relationship among different types of horses were determined by using Bayesian based approach employed in STRUCTURE 2.3.4 software (Pritchard *et al.*, 2000). The analysis was carried out by using the admixture model among race, sport, draft and wild types with burn-in value of 20 000 and 100 000 Markov Chain Monte Carlo (MCMC) repetitions. Runs for each value of K (K = 2–8) were repeated 100 times. The software CLUMP (Jakobsson and A., 2007) and DISTRUCT (Rosenberg, 2004) were used to align 100 replicates and display the results, respectively. The best number of clusters was determined based upon  $\Delta K$  value

**Table 2. Microsatellite marker information.**

Marker	Ch. No	Repeat Motif	Primer sequences	Ta (°C)	Amplicon length (bp)	Fluorescent Dye
AHT4	24q14	(AC)nAT(AC)n	F: AACCGCCTGAGCAAGGAAGT R: CCCAGAGAGTTTACCCT	60	144-164	HEX
AHT5	8	(GT)n	F: ACGGACACATCCCTGCCTGC R: GCAGGCTAAGGAGGCTCAGC	60	126-144	FAM
ASB2	15q21.3-q23	(GT)n	F: CCACTAAGTGTCTTTCAGAAGG R: CACAACCTGAGTTCTCTGATAGG	55	216-250	NED
ASB17	2p14-p15	(AC)n	F: ACCATTCAGGATCTCCACCG R: GAGGGCGGTACCTTTGTACC	60	87-129	NED
VHL20	30	(TG)n	F: CAAGTCCTCTTACTTGAAGACTAG R: AACTCAGGGAGAATCTTCCTCAG	60	87-105	FAM
HMS3	9	(TG)2(CA)2TC(CA)n and (TG)2(CA)2TC(CA)nGA(CA)5	F: CCATCCTCACTTTTTCACCTTTGTT R: CCAACTCTTTGTACATAACAAGA	60	148-170	HEX
HMS6	4	(GT)n	F: GAAGCTGCCAGTATTCAACCATTG R: CTCCATCTTGTGAAGTGTAACCTCA	60	151-169	NED
HMS7	1q25	(AC)2(CA)n	F: TGTGTGTGAAACATACCTTGACTGT R: CAGGAACTCATGTTGATACCATC	60	165-185	FAM
HTG4	9	(TG)nAT(AG)5AAG(GA)5 ACAG(AGGG)3	F: CTATCTCAGTCTTGATTGCAGGAC R: CTCCCTCCCTCCCTCTGTTCTC	55	127-139	HEX
HTG10	21	(TG)n and TATC(TG)n	F: TTTTATTCTGATCTGTACATTT R: CAATTCCC GCCC ACCCCCGCA	55	95-115	HEX

Ch. No. Chromosomal Number; Ta (°C) Annealing temperature

(Evanno *et al.*, 2005) using Structure Harvester application (Earl and Vonholdt, 2011).

## RESULTS

**Genetic diversity:** In the present study ten microsatellite markers (HTG10, ASB2, ASB17, HMS3, HTG04, HMS7, VHL20, AHT05, AHT04 and HMS6) were used to find out the genetic diversity, population structure and phylogenetic relationship among four horse types i.e., draft, race, sports, and wild horses. Summary of genetic variation in terms of observed number of alleles, effective numbers of allele, Shannon's information index, observed heterozygosity,

expected heterozygosity and fixation index for each marker and category is given in table 3 and 4, respectively.

On the basis of 10 microsatellite markers, which were found highly polymorphic for all four types of horses it has been revealed that, a total number of 73 alleles were observed in all studied categories. Number of alleles for each locus ranged from 5 to 11, with an average of 7.3. Maximum number of alleles ( $N_a$ ) were observed for locus ASB17 ( $11.50 \pm 0.19$ ), followed by VHL20 ( $9.50 \pm 0.20$ ), HTG10 ( $8.75 \pm 1.61$ ), ASB2 ( $8.75 \pm 0.25$ ), AHT04 ( $8.25 \pm 0.40$ ), AHT05 ( $8.00 \pm 0.40$ ), HMS6 ( $7.25 \pm 1.31$ ), HMS3 ( $7.00 \pm 0.70$ ), HMS7 ( $6.75 \pm 0.70$ ), and HTG4 ( $5.75 \pm 0.40$ ) as shown in Table 3. Effective number of alleles for each locus ranged from  $3.33 \pm 0.56$  (HTG4) to

**Table 3. Number of different alleles ( $N_a$ ), number of effective alleles ( $N_e$ ), Shannon's information index(I), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), unbiased expected heterozygosity ( $uH_e$ ) and fixation index(F).**

	$N_a$	$N_e$	I	$H_o$	$H_e$	$uH_e$	F
VHL20	9.50(0.20)	5.44(0.44)	1.88(0.04)	0.68(0.0)	0.81(0.01)	0.81(0.01)	0.17(0.07)
HTG4	5.75(0.40)	3.33(0.56)	1.34(0.14)	0.58(0.05)	0.67(0.05)	0.67(0.05)	0.13(0.05)
AHT4	8.25(0.40)	4.36(0.53)	1.66(0.10)	0.69(0.04)	0.76(0.02)	0.76(0.02)	0.09(0.06)
HMS7	6.75(0.70)	4.47(0.45)	1.61(0.08)	0.60(0.07)	0.77(0.02)	0.77(0.02)	0.23(0.10)
AHT5	8.00(0.40)	4.81(0.30)	1.68(0.04)	0.67(0.07)	0.79(0.01)	0.79(0.01)	0.15(0.06)
HMS6	7.25(1.31)	4.48(0.12)	1.65(0.08)	0.70(0.04)	0.78(0.00)	0.78(0.00)	0.14(0.06)
ASB2	8.75(0.25)	4.96(0.52)	1.76(0.08)	0.68(0.10)	0.80(0.02)	0.79(0.02)	0.15(0.10)
HTG10	8.75(1.61)	4.43(1.03)	1.60(0.23)	0.64(0.09)	0.72(0.08)	0.72(0.08)	0.12(0.04)
HMS3	7.00(0.70)	3.94(0.75)	1.50(0.18)	0.65(0.09)	0.70(.08)	0.70(0.08)	0.09(0.05)
ASB17	11.5(0.19)	6.20(0.25)	2.02(0.06)	0.80(0.01)	0.84(0.00)	0.84(0.00)	0.06(0.01)

$N_a$  = No. of Different Alleles;  $N_e$  = No. of Effective Alleles =  $1 / (\sum p_i^2)$ ; I = Shannon's Information Index =  $-1 * \sum (p_i * \ln(p_i))$ ;  $H_o$  = Observed Heterozygosity = No. of Hets / N;  $H_e$  = Expected Heterozygosity =  $1 - \sum p_i^2$ ;  $uH_e$  = Unbiased Expected Heterozygosity =  $(2N / (2N-1)) * H_e$ ; F = Fixation Index =  $(H_e - H_o) / H_e = 1 - (H_o / H_e)$

**Table 4. Number of different alleles (Na), number of effective alleles (Ne), Shannon's information index (I), observed heterozygosity (Ho), expected heterozygosity (He), unbiased expected heterozygosity (uHe) and fixation index (F) within breeds.**

Population	Na	Ne	I	Ho	He	uHe	F
Draft	7.9(0.8)	4.59(0.43)	1.64(0.10)	0.70(0.03)	0.76(0.03)	0.76(0.03)	0.08(0.02)
Race	7.8(0.5)	4.85(0.36)	1.67(0.08)	0.74(0.02)	0.77(0.02)	0.78(0.02)	0.05(0.01)
Sports	9.1(0.6)	5.05(0.33)	1.79(0.06)	0.72(0.02)	0.79(0.01)	0.79(0.01)	0.08(0.02)
Wild	7.8(0.6)	4.08(0.42)	1.57(0.10)	0.49(0.11)	0.71(0.04)	0.72(0.04)	0.31(0.04)

Na = No. of Different Alleles; Ne = No. of Effective Alleles =  $1 / (\sum p_i^2)$ ; I = Shannon's Information Index =  $-1 * \sum (p_i * \ln(p_i))$ ; Ho = Observed Heterozygosity = No. of Hets / N; He = Expected Heterozygosity =  $1 - \sum p_i^2$ ; uHe = Unbiased Expected Heterozygosity =  $(2N / (2N-1)) * He$ ; F = Fixation Index =  $(He - Ho) / He = 1 - (Ho / He)$

$6.20 \pm 0.25$ (ASB17). However, effective number of alleles is less than observed number of alleles, for all markers. The highest value of Na (number of different alleles) was observed for sports horses ( $9.1 \pm 0.6$ ) and respectively, followed by race (Na= $7.8 \pm 0.5$ ), draft (Na= $7.9 \pm 0.8$ ) and Wild horses (Na= $7.8 \pm 0.6$ ) and Likewise, the maximum values of Ne (effective number of alleles) was also observed for sports horses ( $5.05 \pm 0.33$ ) followed by racing (Ne= $4.85 \pm 0.36$ ), draft (Ne= $4.59 \pm 0.43$ ), and Wild horses (Ne= $4.08 \pm 0.42$ ). The highest values of Na and Ne indicate that sports (warm bloods) horses are most diverse category followed by racing, draft, and wild horse category.

Maximum heterozygosity was observed at locus ASB17 ( $0.80 \pm 0.01$ ) having maximum number of alleles ( $11.50 \pm 0.19$ ). Among the 10 microsatellite markers, tested in the present study, ASB17 proved to be highly polymorphic among all categories with highest value of effective number of alleles, observed and expected heterozygosity. Moreover, small difference was seen between observed ( $0.80 \pm 0.01$ ) and expected heterozygosity ( $0.84 \pm 0.00$ ) at locus ASB17. However, HTG4 was found least polymorphic with lowest values of observed (Ho= $0.58 \pm 0.05$ ) and expected heterozygosity (He= $0.67 \pm 0.05$ ). A large difference is found between observed and expected heterozygosity at locus HMS7 and HMS3.

The mean observed heterozygosity value ranged from  $0.49 \pm 0.11$  for wild breeds to  $0.74 \pm 0.02$  for racing breeds. The observed heterozygosity was less than expected heterozygosity for all four categories. The difference between the values of observed and expected heterozygosity were not large for draft, racing, and sports categories. However, in case of wild horses large difference were found between observed ( $0.49 \pm 0.11$ ), and expected heterozygosity ( $0.71 \pm 0.04$ ). Moreover, Unbiased expected heterozygosity for each marker and category is also, separately, given in Table 3 and 4, respectively.

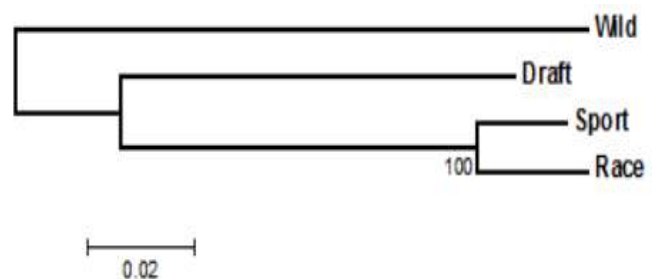
Shannon's index value varied from  $1.57 \pm 0.10$  to  $1.79 \pm 0.06$  in all four horse categories. Its value was highest in Sports breeds, as shown in Table 4, and lowest for Wild breeds of horses. Additionally, the value of Shannon's information index varied for each marker, ranging from  $1.34 \pm 0.14$  to  $2.02 \pm 0.06$ , with maximum value for ASB17 and minimum for

HTG4 locus. Other markers, such as VLH20 ( $1.88 \pm 0.04$ ), ASB2 ( $1.76 \pm 0.08$ ), AHT5 ( $1.68 \pm 0.08$ ), AHT4 ( $1.66 \pm 0.10$ ), HMS6 ( $1.65 \pm 0.08$ ), HMS7 ( $1.61 \pm 0.08$ ) and HTG10 ( $1.60 \pm 0.23$ ) showed moderate value for Shannon diversity index.

Fixation index of each category in relation to total population varied from ( $0.06 \pm 0.01$ ) at ASB17 to ( $0.23 \pm 0.10$ ) at HMS7 locus, as ASB17 showed maximum heterozygosity. Other markers, HMS3 ( $0.09 \pm 0.05$ ), AHT4 ( $0.09 \pm 0.06$ ), HTG10 ( $0.12 \pm 0.04$ ), HTG4 ( $0.13 \pm 0.05$ ), HMS6 ( $0.14 \pm 0.06$ ), ASB2 ( $0.15 \pm 0.10$ ), AHT5 ( $0.15 \pm 0.06$ ) and VLH20 ( $0.17 \pm 0.07$ ), have moderate value of across breeding fixation index.

Moreover, inbreeding coefficient for all four categories varied from  $0.05 \pm 0.01$ , for Racing breeds, to  $0.31 \pm 0.04$  for Wild breeds whereas, moderate values ( $0.080 \pm 0.02$ ) of inbreeding coefficient was observed for draft and sports horse breeds.

**Phylogenetic relationship:** The neighbor joining dendrogram derived from Nei's standard genetic distance of four types of horses is shown in Figure 1 and Table 5. The results of phylogenetic analysis revealed that wild horses were in separate cluster, as predicted. However, draft horses were found much closer to wild horses, whereas race horses were found at maximum far distant from wild horses.

**Figure 1. The neighbor joining dendrogram showing the genetic relationships among the wild, draft, sports and race type horse based on Nei's genetic distance.**

Moreover, it was also observed that though sports horses were in between race, and draft but they were in same cluster with race horses indicating that sports horses are closer to race compared with draft.

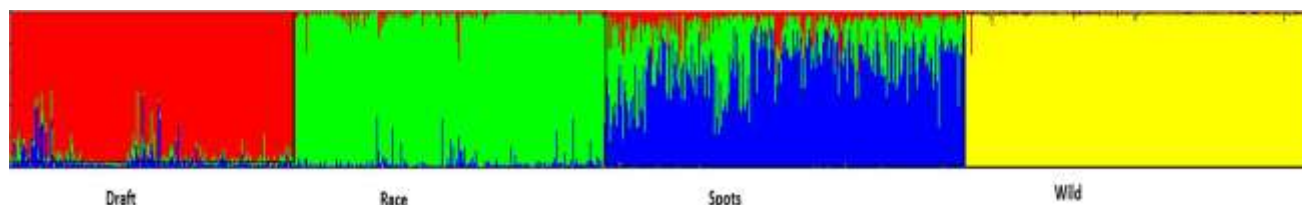


Figure 2. Clustering assignment obtained by STRUCTURE.

Table 5. Nei's genetic distance matrix obtained from allele frequencies of 10 loci in draft, sports, race and wild wild type horses.

	Draft horses	Race horses	Sports horses
Race horses	0.30		
Sports horses	0.27	0.15	
Wild horses	1.37	1.64	1.707

The genetic structure of draft, race, sports and wild types was studied using a model based clustering approach that assigns individuals into one or more types probabilistically based on the allele frequencies detected at different loci. The genetic analysis of these four types of horses was done by using K ranging from 2-7 showed that highest delta K value was at K=4 (Figure 2). When K=4, sports and race types were clustered together and can be regarded as a separate category. At k=5 sports horses separated from other categories and this separation was maintained from k=5 to k=7.

## DISCUSSION

Result indicates that all the makers are highly polymorphic among draft, sports, race and wild type horses. Number of alleles at a locus is a common and easy method to measure genetic diversity within and among populations especially when a population had faced bottleneck. The analysis of data revealed that locus ASB17, is most polymorphic marker among all categories, since has maximum number of alleles (11.50) low polymorphic loci was found HTG4 with (5.75) number of alleles. While sports type have maximum number of alleles for each locus which indicates enough diversity within and among sports breeds of horses. However, maximum heterozygosity was found in race horses.

In this study number of alleles are in accordance with the results of (Khanshour *et al.*, 2013) they found parentage of Arabian and Syrian horses breeds and observed up to 7 number of alleles per locus and maximum number of alleles were for ASB17. Likewise, Berber *et al.* (2014) also observed this highly polymorphic locus with number of alleles 14. All the markers under study were also analyzed by Pires *et al.* (2016) and found ASB17 most polymorphic marker in Nordestino horses while most heterozygous locus was AHT4. Maximum observed heterozygosity was found at ASB17 for all the studied categories, other study reports that AHT4 is

most heterozygous loci among Spanish Celtic breeds of horses (Canon *et al.*, 2001).

However, similar contradiction was also observed in Egyptian horses where HTG10 was the most polymorphic locus (Mahrous *et al.*, 2011). In current study, it was also observed that observed heterozygosity was less than estimated heterozygosity in all four types, which reveals that all the categories exhibit slight deviation from Hardy Weinberg equilibrium. Positive value of inbreeding estimate may also be the reason of this deviation. Among these four categories under study wild type horses showed large deviation from HWE because there is a large difference between observed (0.49) and expected (0.71) heterozygosity.

Heterozygosity of draft breeds (0.71), is comparable with the results of one of the study conducted on heavy draft breeds (Aberle *et al.*, 2004). Mean number of effective alleles, and heterozygosity among draft breeds are also less than racing and sports categories of horses. Results of recent study indicates that draft horse breeds have less diversity as compared to race and sports, because with the development of science and technology, modern means of transportation had replaced the draft breeds of horses. While race and sports breeds are commonly reared for recreation purpose. It was also reported that there is decline in the genetic diversity of coldblooded breeds of horses (Vostry *et al.*, 2011).

In present study, race type showed maximum observed heterozygosity while low heterozygosity was observed in wild type horses which showed contradiction with the results of one of the study where minimum diversity was observed in thoroughbred, which is a race horse breed, when compared this breed with six Indian horse breeds (Gupta *et al.*, 2014). While both studies shows comparable inbreeding coefficient value for racing horses. Value of observed heterozygosity of sports horses is second to race horses, with the value of 0.72. Likewise, Glowatzki-Mullis *et al.* (2006) also observed almost same value of ( $H_o=0.71$ ) for European warm blood horses, which was found highest among all studied breeds.

Shannon's information index is one of the most important and enduring measure to calculate genetic diversity. Value of Shannon's index is moderate for each locus except HTG4, which shows low value. However, low value of Shannon's information index for wild category of horses indicates that, genetic diversity of wild breeds is declining. One of the main reasons of this decline is, inbreeding. Shannon's diversity index is also calculated for each marker in all categories under

study. ASB17 and VLH20 found to be most diverse markers in all types of horses.

Fixation index ( $F_{is}$ ) is one of the important measures used to analyze inbreeding within a breed. Highest value of Fixation index within breed was found in wild horses (0.31), this indicates that high level of inbreeding within wild population and wild breeds suffered loss of genetic variability due to high level of inbreeding, however lowest inbreeding was found for race type which also showed highest value of heterozygosity, overall fixation index within breed was not significantly different from zero for race, draft and sports breeds of horses. This indicates that still sufficient heterozygosity is present within breeds of these categories. Zuccaro *et al.* (2008) also observed positive value of inbreeding coefficient for four studied horse breeds. Results show divergence from study of Avdi and Banos (2008), they also observe genetic variability among endangered horse breeds and found average inbreeding  $0.11(\pm 0.02)$ . Moreover, heterozygosity of wild population also shows deviation from results of Plante *et al.* (2007), they found average heterozygosity 0.626 in wild population, while very low heterozygosity was observed in recent study.

However, across the breed fixation index ( $F$ ) was also calculated for all microsatellite markers. The results indicates that low genetic differentiation at locus ASB17 (0.06), moderate genetic differentiation at locus HMS3 and AHT4, great genetic differentiation was observed at locus HMS7. Overall average value of fixation index indicates that 10% of genetic differentiation can be explained in terms of breed's difference while remaining 90% is due to difference among individuals. The results showed divergence from the results of Behl *et al.* (2007), who reported AHT4 as most polymorphic locus in all breeds under study, with number of alleles 10. they found value of genetic differentiation (0.06) at locus AHT4.

The results of phylogenetic analysis revealed that wild horses were in separate cluster while Draft horses were found much closer to Wild horses, whereas Race horses were found at maximum far distant from Wild horses. Results of the present study are in accordance with Pires *et al.* (2016) who formed a dendrogram of 68 breeds of horse, and reported racing and wild breeds far apart in neighbor joining tree. While Bowling *et al.* (2003) reported that due to an extra chromosome it is difficult to relate directly wild horses with domestic horse breeds. Overall large genetic distance was observed among wild and other three types, one reason of this large genetic distance is an extra pair of chromosome in wild horses (Prezwalski). Secondly, wild horse breeds are endanger, and maintained in captivity. A genetic distance of about 1.707 was observed among sports and wild breeds of horses in this study. Moreover, it was also observed that though Sports horses were in between Race, and draft but they were in same cluster with race horses indicating that sports horses are closer to race compared with draft.

In the current study, minimum genetic distance was observed between race and sports horses. One reason of this minimum genetic distance is that sports horses (warm bloods) are developed by cross between racing (Hot bloods) and draft horses (cold bloods) (Kourkova *et al.*, 2009). While, sports and wild horse breeds are far apart from each other. Large genetic distance between wild and sports horse breeds were observed, which can be explained in terms of limited genetic variation due to hot bloods and cold bloods are ancestors of sports breeds.

The Structure analysis of race, sports, draft and wild types using  $K=2-7$  showed that highest  $\Delta K$  was at  $K=4$ . It showed that sports and racing type horses are clustered together in one cluster while wild and draft breeds are far apart. Current results showed that draft/cold blood horses and sports/warm blood horses are at a distance. Which shows accordance to the results of Leroy *et al.* (2009), who reported that draft and warm blood breeds of horses shows a clear separation. Draft breeds are close wild horses, which was also reported by Khanshour *et al.* (2014) as draft horses are one of the primitive horse category which were domesticated.

**Conclusions:** Currently, breeding of race and sports horse breeds is increasing rapidly in whole world, while wild and draft breeds of horse are endangered. Especially, wild horse breeds had suffered bottle neck several times, which results declining genetic variability among wild breeds. Present study is first of its type in which efforts are made to analyze genetic diversity, population structure and phylogenetic relationship of four categories of horses were studied. We have tried to evaluate genetic diversity as a whole in all types of horses, but due to resources, limited number of markers used in this study. Even then we have stimulating results regarding to genetic diversity and phylogenetic relationship of wild, draft, race and sports category of horses. Since results of present study have shown that there is low number of alleles within all four categories of horses, breeding programs for all kinds of horses should focus on preventing further loss of number of alleles. It is also very compulsory to reduce extensive inbreeding to enhance observed heterozygosity. Moreover, the results can be used for a sustainable breed conservation program.

## REFERENCES

- Aberle, K.S., H. Hamann, C. Drogemuller and O. Distl. 2004. Genetic diversity in German draught horse breeds compared with a group of primitive, riding and wild horses by means of microsatellite DNA markers. *Anim. Genet.* 35:270-277.
- Achmann, R., I. Curik, P. Dovc, T. Kavar, I. Bodo, F. Habe, E. Marti, J. Slkner, and G. Brem. 2004. Microsatellite diversity, population subdivision and gene flow in the Lipizzan horse. *Anim. Genet.* 35:285-292.

- Avdi, M., and G. Banos. 2008. Genetic diversity and inbreeding in the Greek Skyros horse. *Livest. Sci.* 114:362-365.
- Bech, N., C. Novoa, J.F. Allienne and J. Boissier. 2010. Transferability of microsatellite markers among economically and ecologically important galliform birds. *Genet. Mol. Res.* 9:1121-1129.
- Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, F. Bonhomme and I. Montpellier. 1996-2004. GENETIX 4.05, Windows TM software for population genetics. In: *Laboratoire genome, populations, interactions. Montpellier (France):. Universite Montpellier II.*
- Berber, N., S. Gaouar, G. Leroy, S. Kdidi, N. Tabet Aouel and N. Saidi Mehtar. 2014. Molecular characterization and differentiation of five horse breeds raised in Algeria using polymorphic microsatellite markers. *J. Anim. Breed. Genet.* 3:387-394.
- Bigi, D. and G. Perrotta. 2012. Genetic structure and differentiation of the Italian Catria horse. *J. Hered.* 103:134-139.
- Bjornstad, G., N.O. Nilsen and K.H. Roed. 2003. Genetic relationship between Mongolian and Norwegian horses. *Anim. Genet.* 34:55-58.
- Boruszewska, K., M. Lukaszewicz, G. Zieba, A. Witkowski, J. Horbanczuk and K. Jaszcak. 2009. Microsatellite markers may be ineffective in selection of laying hens for polygenic production traits. *Poult. Sci.* 88:932-937.
- Bowling, A.T., W. Zimmermann, O. Ryder, C. Penado, S. Peto, L. Chemnick and T. Zharkikh. 2003. Genetic variation in Przewalski's horses, with special focus on the last wild caught mare. *Cytogenet. Genome. Res.* 102:226-234.
- Canon, J., P. Alexandrino, I. Bessa, C. Carleos, Y. Carretero, S. Dunner, N. Ferran, D. Garcia, J. Jordan, D. Laloë, A. Pereira, A. Sanchez and K. Moazami-Goudarzi. 2001. Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genet. Sel. Evol.* 33:311-332.
- Cheng, H.H. and L.B. Crittenden. 1994. Microsatellite markers for genetic mapping in the chicken. *Poult. Sci.* 73:539-546.
- Cothran, E.G., J.L. Canelon, C. Luis, E. Conant, and R. Juras. 2011. Genetic analysis of the Venezuelan Criollo horse. *Genet. Mol. Res.* 10:2394-2403.
- Diaz, S., F.N. Dulout and P. Peral-Garcia. 2002. Greater genetic variability in Argentine Creole than in Thoroughbred horses based on serum protein polymorphisms. *Genet. Mol. Res.* 1:261-265.
- Earl, D.A. and B.M. vonholdt. 2011. Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359-361.
- Evanno, G., S. Regnaut and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14:2611-2620.
- Felsenstein, J. 1989-2006. PHYLIP (phylogeny inference package). [cited March 10 2013]. Available online at <http://evolution.genetics.washington.edu/phylip.html>.
- Gandini, G.C. and J.K. Oldenbroek. 1999. Choosing the conservation strategy. In: J.K. Oldenbroek (ed.), *Genebanks and the Conservation of Farm Animal Genetic Resources*. DLO, Netherlands. pp.11-13.
- Giacomoni, E.H., G.P. Fernández-Stolz and T.R.O. Freitas. 2008. Genetic diversity in the Pantaneiro horse breed assessed using microsatellite DNA markers. *Genet. Mol. Res.* 7:261-270.
- Glowatzki-Mullis, M.L., J. Muntwyler, W. Pfister, E. Marti, S. Rieder, P.A. Poncet and C. Gaillard. 2006. Genetic diversity among horse populations with a special focus on the Franches-Montagnes breed. *Anim. Genet.* 37:33-39.
- Goodwin, D. 2007. *Horse Behaviour: Evolution, Domestication and Feralisation*. Springer Nature. pp. 1-18.
- Gupta, A.K., M. Chauhan, A. Bhardwaj, N. Gupta, S.C. Gupta, Y. Pal, S.N. Tandon and R.K. Vijh. 2014. Comparative genetic diversity analysis among six Indian breeds and English Thoroughbred horses. *Livest. Sci.* 63:11-17.
- Hanotte, O. and H. Jianlin. 2005. Genetic characterization of livestock populations and its use in conservation decision-making. *Role of Biotechnology, Villa Gualino, Turin, Italy*. pp.131-136.
- Jakobsson, M. and R.N.A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.
- Khanshour, A., E. Conant, R. Juras and E.G. Cothran. 2013. Microsatellite analysis of genetic diversity and population structure of Arabian horse populations. *J. Hered.* 104:386-398.
- Khanshour, A., R. Juras, R. Blackburn and E.G. Cothran. 2014. The legend of the Canadian horse: Genetic diversity and breed. *J. Hered.* 106:37-44.
- Kourkova, L., I. Vrtkova and P. Srubarova. 2009. Microsatellite DNA analysis of genetic diversity in selected horse population. *J. Agrobiol.* 26:57-60.
- Leroy, G., L. Callède, E. Verrier, J.-C. Mériaux, A. Ricard, C. Danchin-Burge and X. Rognon. 2009. Genetic diversity of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genet. Selec. Evo.* 41:1-5.
- Mahrous, K.F., M. Hassanane, M. Abdel Mordy, H.I. Shafey and N. Hassan. 2011. Genetic variations in horse using microsatellite markers. *J. Genet. Eng. Biotechnol.* 9:103-109.
- Miah, G., M.Y. Rafii, M.R. Ismail, A.B. Puteh, H.A. Rahim, S. Ashkani and A. Latif. 2015. Inheritance patterns and

- identification of microsatellite markers linked to the rice blast resistance in BC2F1 population of rice breeding. *Bragantia* 74:33-41.
- Petersen, J.L., J.R. Mickelson, K.D. Cleary and M.E. McCue. 2013. The American Quarter Horse: Population structure and relationship to the Thoroughbred. *J. Hered.* 105:148-162.
- Pires, D.A.F., E.G.A. Coelho, J.B. Melo, D.A.A. Oliverira, M.N. Ribeiro, E.G. Cothran and R. Juras. 2016. Genetic relationship between the Nordestino horse and national and international horse breeds. *Genet. Mol. Res.* 15:1-8.
- Plante, Y., J.L. Vega-Pla, Z. Lucas, D. Colling, B. De March, and F. Buchanan. 2007. Genetic diversity in a feral horse population from Sable Island, Can. *J. Hered.* 98:594-602.
- Pritchard, J.K., M. Stephens and P. Donnelly. 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. Eco. Notes* 4:137-138.
- Sarkar, S., R.L. Pressey, C.R. Margules, T. Fuller and D.M. Stoms. 2006. Biodiversity conservation planning tools: Present status and challenges for the future. *Ann. Rev. Environ. Resour.* 31:123-159.
- Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30:2725-2729.
- Uzans, A.J., Z. Lucas, B.A. McLeod and T.R. Frasier. 2015. Small Ne of the isolated and unmanaged horse population on Sable Island. *J. Hered.* 106:660-665.
- Vicente, M.C.D., F.A. Guzmán, J. Engels and V.R. Rao. 2005. Genetic characterization and its use in decision making for the conservation of crop germplasm. Management. Food and Agriculture organization of the United Nation. 121-128.
- Vila, C. 2001. Widespread origins of domestic horse lineages. *Science* 291:474-477.
- Vostry, L., Z. Capkova, J. Priby, B. Hofmanova, V.H. Vostra and K. Mach. 2011. Population structure of Czech cold-blooded breeds of horses. *Arch. Anim. Breed.* 54:1-9.
- Warmuth, V., A. Eriksson, M.A. Bower, B. G., E. Barrett, B.K. Hanks, S. Li, D. Lomitashvilli, M. Orchir-Goryaeva, V. Sizonov Grigory, V. Soyonov and A. Manica. 2011. Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Pak. Nat. Acad. Sci.* 109:8202-8206.
- Zuccaro, A., S. Bordonaro, A. Criscione, A.M. Guastella, G. Perrotta, M. Blasi and D. Marletta. 2008. Genetic diversity and admixture analysis of Sanfratellano and three other Italian horse breeds assessed by microsatellite markers. *Animal* 2:991-998.