Individual identification and population assignment with microsatellite markers in three Indian donkey populations

RAHUL BEHL*, S K NIRANJAN, JYOTSNA BEHL and R K VIJH

ICAR-National Bureau of Animal Genetic Resources, Karnal, Haryana 132 001 India

ABSTRACT

Ability of the microsatellite markers for individual identification and their potential for breed assignment of individuals was evaluated in three Indian donkey populations. The probability of identity of two random individuals within a population (G1), taking into consideration all eleven loci was 5.91×10^{-13}, 1.06×10^{-12} and 3.67×10^{-14} in Ladakhi, Spiti and Andhra-brown donkeys, respectively. Similarly, the probability of identity of two random individuals from two different populations (G2) between Spiti and Ladakhi donkeys, the two most closely related populations, was only 8.05×10^{-21}. However, the population assignment precision using this set of 11 loci, the correct assignments ranged between 73.08 (Andhra-brown) to 96% (Ladakhi) with frequency method and between 88.46 (Andhra-brown) to 100% (Ladakhi) with Baysian approach. These results suggest that this set of markers can be a promising tool for identification of individuals and their products. Although G2 values were higher than the G1 values but when this set was specifically evaluated for breed allocation purposes, our results indicated that it may require further substantiation before this set can be safely employed for breed/population allocation of individuals in Indian donkey populations.

Keywords: Breed-assignment, Indian-donkeys, Individual identification

Due to their highly polymorphic nature, microsatellite DNA markers have been extensively used for analysis of phylogenetic relationships amongst populations in different species including donkeys (Aranguren-Mendez *et al.* 2001, Jordana *et al.* 2001, Ivankovic *et al.* 2002, Colli *et al.* 2013, Zhang *et al.* 2016, Behl *et al.* 2017a and b, Behl *et al.* 2019). Their usefulness for pedigree or parentage verification has been also evaluated for parentage verification in donkeys (Jordana *et al.* 2001, Aranguren-Mendez *et al.* 2001, Ivanovic *et al.* 2002). Although, their utility in individual demarcation procedures like individual identification and assignment of an individual animal to a breed or population has been evaluated in horses (Bjornstad and Roed 2001, 2002, Kruger *et al.* 2005, Behl *et al.* 2008), no such reports are available in donkeys. A test for the assignment of an individual to a breed is essential for effective and accurate selection/ management of the livestock breeds. Besides, assignment of an individual animal to a population, the discrimination of individual animal is necessary for the authentication of the quality and origin of the livestock products. The present study was undertaken to evaluate a set of eleven microsatellite markers for their potential for individual identification and also to assess their effectiveness in breed assignment of individual animals in three Indian donkey populations.

*Corresponding author e-mail: behl1969@rediffmail.com

MATERIALS AND METHODS

The blood samples were collected from 25 Ladakhi donkeys from the Leh district of Jammu and Kashmir, 46 Spiti donkeys from Lahaul-Spiti district of Himachal Pradesh and 28 Andhra-brown donkeys from Kurnool district of Andhra Pradesh. The DNA was isolated by standard procedure of digestion with proteinase K, extraction with phenol/chloroform and precipitation with ethanol. The stock DNA was stored at –20°C and the working dilutions were stored at 4°C.

The genomic DNA was amplified by PCR using the 11 heterologous microsatellite loci of the horse origin (Aranguren-Mendez 2001, FAO 2011). The details are given in Behl *et al.* 2017b. Each 25 µl reaction consisted of DNA (about 100 ng), primers (7.5 pmol each), dNTPs (200 µM each), 10× buffer (50 mM KCL, 10 Mm tris-HCl, 0.1% gelatin), $MgCl_2$ (1.5 mM) and Taq DNA polymerase (1 unit). The thermocycling conditions included an initial denaturation at 94°C for 2 min followed by 30 cycles of 45 sec at 92°C, 45 sec at annealing temperature (Behl *et al.* 2017b) and 1 min at 72°C. A final extension step was carried out at 72°C for 15 min. The primers were labeled with HEX and FAM to facilitate resolution of alleles on automated DNA sequencer. The allele frequencies and within-breed genetic diversity parameters each locus were calculated using POPGENE computer program version 1.31 (Yeh *et al.* 1999).

The population allocation of individual animals was estimated by phylogenetic approach using Nei's D_A distances (Nei *et al.* 1983) as well as likelihood approach using both frequency method (Paetkau *et al.* 1995) and Bayesian method (Rannala and Mountain 1997) after Monte Carlo resampling (Cornuet *et al.* 1999) with GENECLASS computer package (Piry and Cornuet 1998).

The probability of identity of two random individuals within a population (G1) or from two different populations (G2) was calculated as described by Van-Zeveren *et al.* (1995).

$$GI = \prod_{i=1}^{r}\left[\sum_{j=1}^{ni} q_{ij^4} + 4\sum\sum q_{ij^2}.q_{ik^2}\right]$$

With q_{ij} being the frequency of the jth allele and i^{th} locus in a population.

$$GI = \prod_{i=1}^{r}\left[\sum_{j=1}^{ni} q_{ij^2}.q'_{ij^2} + 4\sum\sum q_{ij}.q'_{ij}.q_{ik}.q'_{ik}\right]$$

where, q and q′ being the frequencies of corresponding alleles.

RESULTS AND DISCUSSION

The within-breed genetic diversity parameters of observed and effective number of alleles, observed and expected heterozygosity at each locus in each population and genetic distances between these populations were published in our earlier publications (Behl *et al.* 2017a and b, Behl *et al.* 2019). The probability of identity of two random individuals within a population (G1), taking into consideration all eleven loci varied between 3.67×10^{-14} (Andhra-brown) to 1.06×10^{-12} (Spiti) (Table 1). The probability of identity of two random individuals from two different populations (G2) with these eleven loci was 8.05×10^{-21} between Ladakhi and Spiti donkeys (Table 2). Due to the absence of even a single common allele at some of the loci (HTG6, AHT4, NVHEQ54, COR71), the G2 between Spiti and Andhra-brown donkeys was zero at these

Table 1. Probability of identity of two individuals chosen at random from within a population (G1) at 11 microsatellite loci in three populations of Indian donkeys

Microsatellite locus	G1		
	Ladakhi	Spiti	Andhra-Brown
HTG15	0.174	0.060	0.062
HTG7	0.033	0.023	0.008
HTG10	0.166	0.096	0.070
HMS2	0.031	0.152	0.017
COR18	0.042	0.033	0.031
VHL209	0.401	0.257	0.381
AHT4	0.037	0.029	0.021
NVHEQ54	0.040	0.167	0.132
COR7	0.095	0.203	0.144
HTG6	0.111	0.118	0.088
COR71	0.076	0.055	0.141
Cumulative G1	5.91×10^{-13}	1.06×10^{-12}	3.67×10^{-14}

Table 2. Probability of identity of two individuals chosen at random from two different populations (G2) at 11 microsatellite loci in three populations of Indian donkeys

Microsatellite locus	G2		
	Ladakhi and Spiti	Ladakhi and Andhra-brown	Spiti and Andhra-brown
HTG15	0.091	0.008	0.004
HTG7	2×10^{-4}	8×10^{-4}	0.001
HTG10	0.097	0.065	0.052
HMS2	0.010	0.010	0.025
COR18	0.003	0.027	0.002
VHL209	0.308	0.39	0.301
AHT4	0.002	0.023	0.000
NVHEQ54	0.002	0.032	0.000
COR7	0.130	0.000	3.03×10^{-7}
HTG6	0.113	0.000	0.000
COR71	0.010	0.000	0.000
Cumulative G2 at loci with G2 of >0	8.05×10^{-21} (all 11 loci)	1.05×10^{-18} (7 loci)	3.02×10^{-14} (8 loci)
Cumulative G2 at all 11 loci	8.05×10^{-21}	0.000	0.000

loci resulting in cumulative G2 at all the eleven loci to be zero. At other seven loci, the cumulative G2 was 1.05×10^{-18}. Similarly, G2 was zero at three loci (HTG6, COR7, COR71) between Ladakhi and Andhra-brown donkeys due to absence of even a single common allele. The cumulative G2 at other eight loci between Ladakhi and Andhra-brown donkeys was 3.02×10^{-14}. These values were clearly lower than the G1 values discussed above indicating that the probability of identity of two random individuals was clearly less between two individuals from different populations than from within a population. These values also showed the suitability of these loci to distinguish individual donkeys or their products from two different populations or within a population. The cumulative G2 even with only six loci (HTG15, HTG7, HTG10, HMS2, COR18, VHL209) was as low as 1.76×10^{-11}, 3.47×10^{-12} and 4.12×10^{-11} between Ladakhi/ Spiti, Spiti/ Andhra-brown and Ladakhi/ Andhra-brown donkeys, respectively. The cumulative G1 at these six loci were 5.03×10^{-7}, 1.66×10^{-7} and 7.35×10^{-9} in Ladakhi, Spiti and Andhra-brown donkeys, respectively.

Besides distinguishing between individuals in breeding/ conservation programmes, the allocation of an individual to a population is equally important to discriminate between pure-breds and cross-breds for skilful management of the animal genetic resources. If a method could be developed for authentication of breed or population of an individual it could be of great help to the breeders. Although, the possibilities of using microsatellites for assigning breed identities to anonymous samples have been evaluated in some horse breeds (Canon *et al.* 2000, Bjornstad and Roed 2001, 2002, Behl *et al.* 2008), no such reports are available in donkeys. We attempted to evaluate the potential of the above set of 11 microsatellite loci for population assignment in Indian donkey populations and breeds using phylogenetic

Table 3. Percent of unambiguously assigned animals of three Indian donkey populations after allocation with phylogenetic approach using Nei's D_A distances and likelihood analysis both with frequency method and Bayesian method after Monte Carlo resampling using 11 microsatellite loci

Breed	Percent assignment precision		
	Phylogenetic approach	Frequency method	Bayesian method
Ladakhi	100.00	96.00	100.00
Spiti	100.00	93.48	93.47
Andhra-brown	96.15	73.08	88.46

approach as well as likelihood approach using both frequency method (Paetkau *et al.* 1995) and Bayesian method (Rannala and Mountain 1997).

In phylogenetic approach with Nei's D_A distances (Nei *et al.* 1983) using 11 microsatellite loci, all the individuals could be unambiguously assigned to the respective population in Ladakhi and Spiti donkeys (Table 3). In the case of Andhra-brown donkeys, assignment accuracy of 96.15% was observed (Table 3). Only those animals were considered to be unambiguously assigned to a population that had an assignment probability to that population to be clearly higher than that of the second most probable population. If the ratio of the most likely allocation with the second most likely allocation approaches one, it is assumed that there is ambiguity in the assignment of the particular animal (Banks and Eichert 2000).

Similarly, in the likelihood approach, a population is almost always designated because there is always a most likely or a closest population in a reference set. In practical conditions the animal to be assigned may not belong to any of the populations under consideration (Cornuet *et al.* 1999). Therefore, a minimum assignment probability of 0.05 was adopted for unambiguous assignment of an individual to a population. With this stringency, with frequency method (Paetkau *et al.* 1995), with this set of eleven loci, the correct assignments ranged between 73.08 (Andhra-brown) to 96.00% (Ladakhi). With Baysian method (Rannala and Mountain 1997), the correct assignments ranged between 88.46 (Andhra-brown) to 100% (Ladakhi). Although no reports are available for donkey populations, similar levels of population assignment precision were reported by Tozaki *et al.* (2003) in Japanese and Asian horse breeds. Similar observations were also made earlier in Indian horse breeds also that included the Marwari breed which is considered to be comparatively a more pure-bred stock (Behl *et al.* 2008).

The comparatively lower assignment precision obtained in our study for breed assignment with microsatellites in Indian donkey populations are in disagreement to the earlier similar studies that have proposed the efficacy of microsatellite loci for assigning breed identities to anonymous equine samples based on their studies in Spanish Celtic and Norwegian horses (Canon *et al.* 2000, Bjornstad and Roed 2001, 2002). Several factors have been proposed to affect the accuracy of such individual specific demarcation procedures such as genetic differentiation between the populations in question and degree of reproductive isolation etc. (Cornuet *et al.* 1999). One possible reason for comparatively lower assignment precision in our study could be that the livestock breeds and populations including the donkeys in developing countries like India may not be as well differentiated and purebred stock as the Western breeds.

The above results suggest that although this set of eleven microsatellite loci can be safely employed for identification of individuals and their products both form within a population or from two different populations. Although, cumulative G2 values were lower than the G1 values but when this set was specifically evaluated for assignment of an individual to a breed/population, our results indicated that it may require further substantiation before they can be safely employed for breed allocation purpose in Indian donkeys.

The results of this study suggest that this set of markers can be a promising tool for identification of individual donkeys and their products. However, in case of allotment of an individual donkey to a population, our results indicated that it may require further substantiation before this set can be safely employed for breed/population allocation of individuals.

REFERENCES

Aranguren-Mendez J, Jordana J and Gomez M. 2001. Genetic diversity in Spanish donkey breeds using microsatellite DNA markers. *Genetics Selection Evolution* **33**: 433–42.

Banks M A and Eichert W. 2000. Whichrun–a computer program for population assignment of individuals based on multilocus genotype data. *Journal of Heredity* **91**: 87–89.

Behl R, Behl J, Gupta N and Gupta S C. 2008. Evaluation of microsatellite genotyping based individual assignment in five Indian horse breeds. *Indian Journal of Animal Sciences* **78**: 384–87.

Behl R, Niranjan S K, Behl J, Sharma H, Tantia M S, Arora R, Ranjan P, Bharati V K, Iqbal M, Vijh R K and Sharma A. 2019. Genetic characterization of Ladakhi donkeys using microsatellite markers. *Indian Journal of Animal Sciences* **89**: 157–60.

Behl R, Niranjan S K, Behl J, Tantia M S, Arora R, Dharma-Rao M V, Reddy P P, Vijh R K and Sharma A. 2017a. Genetic characterization of Brown type donkeys of Andhra Pradesh. *Indian Journal of Animal Sciences* **87**: 1102–05.

Behl R, Sadana D K, Behl J, Banerjee P, Joshi J, Vijh R K, Attri P N, Nadda S and Joshi B K. 2017b. Characterization and

microsatellite analysis for genetic diversity and bottlenecks of Spiti donkeys. *Indian Journal of Animal Sciences* **87**: 1221–25.

Bjornstad G and Roed K H. 2001. Breed demarcation and potential for breed allocation of horses assessed by microsatellite markers. *Animal Genetics* **32**: 59–65.

Bjornstad G and Roed K H. 2002. Evaluation of factors affecting individual assignment precision using microsatellite data from horse breeds and simulated breed crosses. *Animal Genetics* **33**: 264–70.

Colli P G, Negrini R, Bomba L, Bigi D, Zambonelli P, Verini S A, Liotta L and Ajmore-Marron P. 2013. Detecting population structure and recent demographic history in endangered livestock breeds: the case of the Italian autochthonous donkeys. *Animal Genetics* **44**: 69–78.

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

Cornuet J M, Piry S, Luikart G, Estoup A and Solignac M. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**: 1989–2009.

FAO. 2011. ISAG-FAO recommended microsatellite markers – Donkey (Appendix 7). *FAO animal production and health guidelines* (*No 9*)—*Molecular genetic characterization of animal genetic resources*, FAO Publications, Rome. pp. 78–79.

Ivankovic A, Kavar T, Caput P, Mioc B, Pavic V and Dovc P. 2002. Genetic diversity of three donkey populations in the Croatian coastal region. *Animal Genetics* **33**: 169–77.

Jordana J, Folch P and Aranguren J A. 2001. Microsatellite analysis of genetic diversity in Catalonian donkey breeds. *Journal of Animal Breeding and Genetics* **118**: 57–63.

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

Nei M, Tajima F and Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**: 153–70.

Paetkau D, Calvert W, Stirling I and Strobeck C. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**: 347–54.

Piry S and Cornuet J M. 1998. GENECLASS, A program for assignation and exclusion using molecular markers, available at http: //www.ensam.inra.fr/CBGP. Campus International de Baillarguet, Cedex, France.

Rannala B and Mountain J L. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of National Academy of Sciences USA* **94**: 9197–221.

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

Van-Zeveran A, Peelman L, Van de Weghe A and Bouquet Y. 1995. A genetic study on four Belgian pig populations by means of seven microsatellite loci. *Journal of Animal Breeding and Genetics* **112**: 191–204.

Yeh F C, Boyle T, Rongcal Y, Ye Z and Xian J M. 1999. Popgene, version 3.31, a Microsoft Windows based freeware for population genetic analysis (https://sites.ualberta.ca/~fyeh/popgene.html).

Zhang R F, Xie W M, Zhang T and Lei C Z. 2016. High polymorphism at microsatellite loci in the Chinese donkey. *Genetics and Molecular Research* **15**: gmr15028291.