Morphometric and microsatellite-based comparative genetic diversity analysis in *Bubalus bubalis* from North India

Vikas Vohra [Corresp., 1] , **Narendra P Singh** [1] , **Supriya Chhotaray** [1] , **Varinder S Raina** [2] , **Alka Chopra** [3] , **Ranjit S Kataria** [4]

[1] Animal Genetics and Breeding Division, National Dairy Research Institute, Karnal, Haryana, India

[2] Department of Animal Husbandry and Dairying, Ministry of Fisheries, Animal Husbandry and Dairying, New Delhi, New Delhi, India

[3] Animal Biotechnology Centre, National Dairy Research Institute, Karnal, Haryana, India

[4] Animal Biotechnology Division, ICAR – National Bureau of Animal Genetic Resources, Karnal, Haryana, India

Corresponding Author: Vikas Vohra
Email address: Vikas.Vohra@icar.gov.in

To understand the similarities and dissimilarities of a breed structure among different buffalo breeds of North India, it is essential to capture their morphometric variation, genetic diversity, and effective population size. In the present study diversity among three important breeds, namely, Murrah, Nili-Ravi, and Gojri were studied using a parallel approach of morphometric characterization and molecular diversity. Morphology was characterized using 13 biometric traits, and molecular diversity through a panel of 22 microsatellite DNA markers recommended by FAO, Advisory Group on Animal Genetic Diversity, for diversity studies in buffaloes. Canonical discriminate analysis of biometric traits revealed different clusters suggesting distinct genetic entities among the three studied populations. Analysis of molecular variance revealed 81.8% of genetic variance was found within breeds, while 18.8% of the genetic variation was found between breeds. Effective population sizes estimated based on linkage disequilibrium were 142, 75, and 556 in Gojri, Nili-Ravi, and Murrah populations, respectively, indicated the presence of sufficient genetic variation and absence of intense selection among 3 breeds. The Bayesian approach of STRUCTURE analysis (at K=3) assigned all populations into 3 clusters with a degree of genetic admixture in the Murrah and Nili-Ravi buffalo populations. Admixture analysis reveals that the Murrah and Nili-Ravi while identified the Gojri as unique buffalo germplasm, indicating that there might be a common origin. The study provides important insights on buffalo breeds of North India that could be utilized in designing an effective breeding strategy, with an appropriate choice of breeds for upgrading local non-descript buffaloes along with conservation of unique germplasm.

1   **Morphometric and microsatellite-based comparative genetic diversity analysis in *Bubalus***
2   ***bubalis* from North India**

4   **Vikas Vohra[1][†], Narendra Pratap Singh[1], Supriya Chhotaray[1], Varinder Singh Raina[2], Alka**
5   **Chopra[3], and R.S. Kataria[4]**

7   [1]Animal Genetics Breeding Division, ICAR-National Dairy Research Institute, Karnal-132001,
8   Haryana, India.

9   [2]Consultant, Department of Animal Husbandry and Dairying, Ministry of Fisheries, Animal
10   Husbandry and Dairying, New Delhi.

11   [3]Animal Biotechnology Centre, ICAR-National Dairy Research Institute, Karnal-132001,
12   Haryana, India

13   [4]ICAR – National Bureau of Animal Genetic Resources, Karnal-132001, Haryana, India.

14   [†]**Corresponding author:**

15   Dr. Vikas Vohra, Principal Scientist

16   Address: Animal Genetics Breeding Division, ICAR-National Dairy Research Institute, Karnal-
17   132001, Haryana, India

18   Email: Vikas.Vohra@icar.gov.in

19   vohravikas@gmail.com

20   **Abstract**

21   To understand the similarities and dissimilarities of a breed structure among different buffalo

22   breeds of North India, it is essential to capture their morphometric variation, genetic diversity, and

23   effective population size. In the present study diversity among three important breeds, namely,

24   Murrah, Nili-Ravi, and Gojri were studied using a parallel approach of morphometric

25   characterization and molecular diversity. Morphology was characterized using 13 biometric traits,

26   and molecular diversity through a panel of 22 microsatellite DNA markers recommended by FAO,

27   Advisory Group on Animal Genetic Diversity, for diversity studies in buffaloes. Canonical

28   discriminate analysis of biometric traits revealed different clusters suggesting distinct genetic

29   entities among the three studied populations. Analysis of molecular variance revealed 81.8% of

30   genetic variance was found within breeds, while 18.8% of the genetic variation was found between

31   breeds. Effective population sizes estimated based on linkage disequilibrium were 142, 75, and

32   556 in Gojri, Nili-Ravi, and Murrah populations, respectively, indicated the presence of sufficient

33   genetic variation and absence of intense selection among 3 breeds. The Bayesian approach of

34   STRUCTURE analysis (at K=3) assigned all populations into 3 clusters with a degree of genetic

35   admixture in the Murrah and Nili-Ravi buffalo populations. Admixture analysis reveals that the

36   Murrah and Nili-Ravi while identified the Gojri as unique buffalo germplasm, indicating that there

37   might be a common origin. The study provides important insights on buffalo breeds of North India

38   that could be utilized in designing an effective breeding strategy, with an appropriate choice of

39   breeds for upgrading local non-descript buffaloes along with conservation of unique germplasm.

40   **Keywords: Buffalo, Genetic Variation, India, Microsatellite Markers**

42 **Introduction**

43 Through the course of evolution, forces such as mutation, adaptation, reproductive isolation,

44 random drift, selection and breeding have created vast diversity (breeds) among the Bubalus

45 bubalis in India. Many well-defined breeds have been formed for various purposes and the high-

46 performance breeds are intensely selected worldwide. This has led to replacement of native low

47 performance breeds with high performance ones causing erosion of genetic resources (Groeneveld

48 et al., 2010). However, future trait of interests concerns the genetic diversity of low-performance

49 breeds so, that the population can be maintained for future breeding (Notter, 1999). The genetic

50 diversity of livestock species such as cattle, buffalo, pig, sheep, goat, camel and chicken have been

51 widely studied in several population and the diversity of zebu and taurine cattle breeds is one of

52 the most studied. Diversity between swamp and riverine buffaloes also have been studied using

53 microsatellite markers, mitochondrial D-loop and cytochrome b sequence variations (Barker et al.,

54 1997; Lau et al., 1998; Zhang et al., 2007).

55 Over the past few decades, world buffalo population has rapidly increased with 208 million

56 buffaloes in the world presently, having 96.79% of population in Asia, 1.68% in Africa, 1.23% in

57 the Americas and 0.22% in Europe (FAOSTAT, 2019). India alone has 109.85 million of buffalo

58 population with 17 registered breeds (20th All India Livestock Census, 2019). Water buffalo

59 (Bubalus Bubalis) that most probably domesticated in Indus Valley region for multiple utility

60 creates a rich Bubaline diversity in Northern regions of India, comprising states of Punjab,

61 Haryana, Himachal Pradesh, Delhi and Western Uttar Pradesh. The predominant bubaline genetic

62 resources documented from the region include Murrah, Nili Ravi and Gojri buffaloes

63 (http://www.nbagr.res.in/regbuf.html). Murrah being dominating buffalo germplasm with superior

64 milk-producing ability has suppressed the need for identification and characterization of other

65   breeds. On the other hand, Gojri is one of the little-known buffalo population of the region, with a

66   good milch potential on low to zero input system of dairying and is maintained on a semi-migratory

67   extensive system of management (Vohra et al., 2012 and 2015).

68   Characterization and classification of animal genetic resources (AnGR) require ample knowledge

69   of the geographical distribution of the breeds, identification of unique characteristics, population

70   size and structure, production environment, and genetic diversity. It is customary to perform a

71   detailed molecular study along with physical and phenotypic assessment to check within and

72   between population diversity in order to characterize a population (Weitzman, 1993; Hall &

73   Bradley, 1995; Barker, 1999; Ruane, 2000; Bruford et al., 2003; Simianer, 2005; Toro & Caballero,

74   2005). Vohra et al. (2015) have used 13 morphometric traits of Gojri buffaloes for phenotypic

75   characterization using Principal Component Analysis, a multivariate statistical technique.

76   Multivariate statistical analysis techniques viz. classical principal component analysis serves the

77   objectives of dimension reduction and clustering when multiple morphometric traits are measured

78   (Johanson and Wichern, 2002; Yadav et al., 2017).

79   The neutrality, co-dominant inheritance and high polymorphic information content of

80   microsatellite markers have rendered them as the markers of choice for diversity studies (Metta et

81   al., 2004; Li et al., 2005; Yoon et al., 2005; Sodhi et al., 2005, 2006; Kumar et al., 2006; Pandey

82   et al., 2006a; Pandey et al., 2006b; Vijh et al., 2008; Sharma et al., 2013). Several genetic diversity

83   studies of water buffalo populations have been carried out throughout the world using

84   microsatellite markers (Kataria et al., 2009; Ángel-Marín et al., 2010; Gargani et al., 2010;

85   Mekkawy et al., 2012; Ünal et al., 2014; Vohra et al., 2017; Khade et al., 2019).

86   The genetic diversity within Murrah, Nili-Ravi and Gojri breeds have been studied independently

87   that share the common breeding tract in North India. However, in India, buffalo breeding is largely

88   restricted to natural mating that subsequently may have led to admixture of these populations.

89   Hence, there is a need to assess the between-breed genetic diversity among these breeds. The

90   present study was performed to assess the levels of genetic diversity, and population structure

91   among three buffalo breeds of North India. The results will help in formulating an effective

92   breeding, management policy, shaping future conservation plans for maintaining breed purity and

93   reducing the possible admixture due to introgression among purebreds. Thus, it is imperative to

94   compare the region-specific diversity and breed status of bubaline germplasm.

95   **Materials & Methods**

96   **Sampling strategy**

97   Sampling was done from their respective native tracts, to compare the genetic diversity between

98   three different breeds. Gojri buffalo samples were collected during 2017-18 from areas of Punjab

99   and Himachal Pradesh ($30^0\,9'$ to $32^0\,3'$ N and 75^0 to 77^0 E) states of India, and samples for Nili-

100   Ravi buffaloes were collected from Punjab state (28° 17' to 32° 17' N and 74° to 76° 41' E). The

101   Nili-Ravi has a comparatively smaller geographical distribution compared to Murrah and Gojri. In

102   India, Murrah buffaloes are found in almost all regions but its native area is Haryana state ($28^0\,02'$

103   to $30^0\,21'$ N and 75^0 to 77^0 E) hence, sampling was performed from Haryana and Punjab. The data

104   of Murrah and Nili-Ravi was taken for comparative analysis from Buffalo Genomics Lab of

105   National Bureau of Animal Genetic Resources, Karnal. The breeding and sampling tract had a herd

106   size of 2-6 buffaloes per households. To ensure that selected animals are unrelated, in the absence

107   of detailed pedigree accounts, buffalo breeders were interviewed in detail and their records were

108   checked. Only those animals who were not having common parents for at least 3-4 generations

109   were included in the study. Buffaloes were selected for this study following guidelines of

110   measurement of domestic animal diversity program (FAO, 2011) those represented the original

111   indigenous true to type phenotype. Blood samples were collected with the consent of herd owners.

112   Approximately 5-10 ml of blood from jugular vein was collected by trained Veterinarian using

113   aseptic measures. All the studies were carried out under approval of ICAR-National Dairy

114   Research Institute IAEC 1705/GO/ac/13/CPCSEA.

115   Morphometric traits were measured on a total of 242 adult female buffaloes, comprising of 113

116   Murrah, 37 Nili-Ravi, and 92 Gojri buffaloes, to avoid the sex and age differences. Thirteen (13)

117   different traits were measured on all three breeds as suggested by Breno et al (2018). All the

118   measurements on the animal were recorded in their normal standing position on a levelled surface

119   using a tape measure by the same technical person. Traits recorded were body height (HT), body

120   length (BL), chest girth (CG), paunch girth (PG), face length (FL), face width (FW), horn length

121   (HL), horn circumference (HC), ear length (EL), distance between hip bone (HB), distance

122   between pin bone (PB), tail length (TL), and tail length up to switch (TS). To avoid age effects,

123   only adult buffaloes (3.5 years above) were included in study. For microsatellite genotyping, blood

124   samples were collected from 128 (40 Murrah, 40 Nili-Ravi, and 48 Gojri) buffaloes.

125   **Genotyping microsatellite markers**

126   Genomic DNA was isolated from blood samples by standard phenol–chloroform extraction

127   protocol, as described by Sambrook and Russel (2001). DNA concentration was checked by

128   spectrophotometric method. Genetic variation was assayed using 25 microsatellite markers.

129   Microsatellite genotyping was carried out as previously describe in Vohra et al. (2017) following

130   the protocol of Mishra et al. (2010). Fluorescent-tagged forward primers for each microsatellite

131   were used. The primers those were able to produce a fragment size >75bp were used in the study.

132   Fragment length analysis was performed through ABI PRISM 3100 automatic sequencer (Applied

133   Biosystems, Foster City, CA, USA) after performing polymerase chain reaction (PCR) for

134   fragment amplification. Allele length for the different fragments generated was determined as

135   described in Vohra et al. (2017) using GeneScan software (version 5.0 Applied Bio system).

136   Observed number of alleles (N_a), theta estimate (θ_H), expected heterozygosity (H_e), F_{IT} (total

137   inbreeding estimate), F_{ST} (measurement of population differentiation) and F_{IS} (within- population-

138   inbreeding estimate) were calculated using Arlequin v3.5 (Excoffier et al., 2010). Pairwise

139   differences between populations using molecular distances were calculated. Molecular diversity

140   indices were calculated as per Tajima (1983), Tajima (1993), Nei (1987), and Zouros (1979),

141   implemented in Arlequin v3.5, and allowing 5% level of missing data. Analysis of molecular

142   variances was done using 1,000 permutations. Exact test of population differentiation was

143   performed with 1,00,000 Markov chain steps and 10,000 dememorization steps.

144   **Statistical analysis**

145   Statistical analyses on morphometric data were performed using SPSS v17.0 software (SPSS,

146   2001). Multivariate analysis technique such as canonical discriminant analysis (CDA)

147   simultaneously analyse multiple correlated measurements in a single individual and increases the

148   discriminatory power by eliminating variables explaining less variation in the dataset. The relation

149   between the group the individual belongs to and a set of morphometric traits are quantified using

150   CDA (Zhao and Maclean, 2000). As, CDA provides optimum discrimination between population

151   to classify them as a different breed hence, widely used in breed characterization and genetic

152   diversity studies.

153   The canonical discriminant analysis was performed in SAS v9.3 program (SAS Institute Inc., 2011)

154   using Proc disc procedure, for determining the most discriminatory morphometric traits. The

155   probabilities of assigning an individual to a population were determined using Discrim procedure

156   based on the linear discriminant function that included the thirteen morphometric variables. Wilk's

157   Lambda was used as the test statistics to check for the differences between the means of identified

158   groups of subjects on a combination of dependent variables.

159   Population assignment was performed using the Bayesian Markov chain Monte Carlo approach

160   implemented in Structure v2.3.4 (Pritchard et al., 2000). The Bayesian clustering algorithm

161   simultaneously estimates allele frequencies at each and individuals are assigned probabilistically

162   to one of the K subpopulations. It assumes that prior distribution of population to which individuals

163   belong and allele frequencies are known.

164   The most likely number of subpopulations was determined by the Evanno ΔK method (Evanno et

165   al., 2005) using R package "POPHELPER" (Francis et al., 2017). Twenty independent runs were

166   performed for $K = 2$ to 4 to identify the most likely number of clusters present in the dataset. The

167   analysis was performed with a burn in period of 10000 and 50000 MCMC iterations. Effective

168   population size (N_e) was checked for the three population. N_e was estimated using linkage

169   disequilibrium method using NeEstimator v2.01 (Do et al., 2014) Software. The P-critical value

170   (rare allele frequency) was set to 0.05, below which all the alleles were rejected. Jackknife

171   confidence intervals (CI) were calculated for each estimate, N_e, of different population.

172   Discrimination between populations was elucidated graphically through principal coordinate

173   analysis (PCoA) using Darwin v6.0.021 (Perrier et al., 2003). Principal coordinate analysis is a

174   classical multidimensional scaling method based on dissimilarity or distance matrix to assign each

175   individual a location in a two or three-dimensional space. The dissimilarity matrix based

176   phylogenetic tree was also obtained through Darwin.

177   **Results**

178   **Classificatory analysis based on Morphometric traits**

179   The means and standard deviation, coefficient of variations and comparison of mean difference

180   between populations for each trait across population is listed in Table 1. A Canonical Discriminant

181   analysis was used to compare different morphometric traits and first two canonical discriminant

182   functions were used in the analysis, which explained 66.7% and 33.3% of total variance,

183   respectively. Wilk's Lambda was used as the test statistics to check the difference between means

184   of the two groups and was found to be significant (Table 2). Classification based on canonical

185   discriminant functions for both original and cross-validated counts predicted 100% assignment of

186   each adult buffaloes to their hypothetically known populations *i.e.* Murrah, Nili-Ravi, and Gojri.

187   All the individuals plotted based on 1st and 2nd canonical discriminant functions were clustered into

188   three distinct groups suggesting three different breeds in the sample (Fig. 1).

189   **Microsatellite variations**

190   Among 25 microsatellite loci genotyped for this study, only 22 loci that were polymorphic for all

191   three populations were used for further downstream analysis. A total of 145, 138, and 173 alleles

192   were found across 22 loci in the 128 individuals sampled from the Murrah, Nili-Ravi, and Gojri

193   buffaloes, respectively. ILSTS60 was highly polymorphic in Gojri buffaloes, ILSTS95 in both

194   Murrah and Nili-Ravi and ILSTS61 in Murrah (Fig. S1a). Mean number of alleles for all

195   populations varied from 3.67±2.08 at ILSTS19 to 10.33±0.58 at CSSM47. Mean expected

196   heterozygosity (H_e) across all populations ranged from 0.14±0.02 at ILSTS19 to 0.81±0.04 at

197   ILSTS58. The mean H_e estimated over all loci was lowest in Murrah (0.58±0.25) while it was

198   highest in Gojri population (0.70±0.15) (Fig. S1b). Estimator of mutation parameter (θ_H) that is

199   obtained using observed homozygosity values was estimated under infinite allele model. Mean θ_H

200   ranged from 1.36 in Murrah to 2.33 in Gojri buffaloes (Fig. S1c). Across all three populations mean

201   θ_H ranged from 0.17±0.03 (ILSTS19) to 4.44±1.16 (ILSTS58). Marker wise number of alleles, H_e,

202   and θ_H in each breed given in Table 3.

203   **Genetic diversity**

204   Global Analysis of molecular variance (AMOVA) using 19 polymorphic loci was accomplished.

205   Wright's F-statistics values obtained from the results of global AMOVA revealed 11.7% deficit of

206   heterozygotes for each of the analyzed breeds (F_{IS}) whereas the total population had a 27.8% deficit

207   of heterozygotes (F_{IT}). The average genetic differentiation (F_{ST}) between the breeds was 18.2% (p

208   = 0.00001) indicating significantly higher discrimination between breeds (Table 4). Details of

209   AMOVA results are presented in Table 4. The pair-wise F_{ST}, Slatkin linearized F_{ST}, and Nei's

210   distance (d) values were used to illustrate the genetic distance between breeds (Fig. 2a, 2b & 2c),

211   which significantly differentiated all three breeds.

212   Murrah and Nili-Ravi population were clustered together while the Gojri population was present

213   as a distinct group, suggesting it as a different breed in factorial correspondence analysis (Fig. 3)

214   and phylogenetic tree (Fig. S2).

215   Effective population size (N_e) was estimated excluding rare alleles with an allele frequency below

216   0.05. The estimated effective population size of Gojri, Nili-Ravi, and, Murrah was found to be 142,

217   75, and 556, respectively. The Jack-knife CIs for the Ne estimates were 83-396, 48-141, and 136

218   to infinity for Gojri, Nili-Ravi, and Murrah, respectively at 0.05 P-critical value of rare alleles.

Bayesian genetic structure

220 Number of possible sub-populations estimated through Evanno ΔK method suggested a maximum

221 of three populations (Fig. 4). Population assignment accomplished in STRUCTURE for K=2, 3,

222 and 4 and results are presented in the form of bar plot (Fig. 5). For K=3, as estimated through

223 Evanno ΔK method, it showed 99.4% of Gojri buffaloes are classified into their pre-defined breed.

224 95.9% of Nili-Ravi and 83.6% of Murrah were assigned to their respective pre-defined groups.

225 Inferred ancestry of each individual (for K=3) along with average proportion of each individuals

226 classified into respective pre-assigned breeds (for K=2, 3, and 4) is reported (Table 5).

Discussion

228 In India, limited work on complete characterization and classification of buffalo genetic resources

229 have been carried out in past, primarily due to availability of much acclaimed Murrah buffaloes.

230 The native breeding tract of Murrah buffalo is North India, and currently, more than 40% of the

231 countries buffalo population is either Murrah or has been crossed with Murrah buffaloes. Hence,

232 genetic studies on other buffalo populations is often neglected. However, several studies have been

233 taken up on morphometric characterization of individual breeds yet there are limited reports on

234 genetic diversity studies through molecular markers and comparative studies. A parallel approach

235 of characterization and classification of buffalo germplasm in a region is much needed for genetic

236 improvement in such populations. The present study is an evaluation of Riverine buffaloes of North

237 India taking a geographical region-based approach.

Morphological diversity

239 Gojri animals with unique phenotypic appearance are quite distinct from Murrah, Murrah crosses,

240 and Nili Ravi (Vohra et al., 2012). The average measurements for body biometric traits across the

241   studied buffalo populations of the North India is listed in Table 1. Thirteen body biometric traits

242   across 3 population when compared, revealed significant differences among the studied

243   populations, except for FL and EL among Murrah and Gojri buffaloes, HC and HL between Nili-

244   Ravi and Gojri population. Body height (HT) and face width (FW) did not vary among Murrah and

245   Nili-Ravi populations. The comparison of morphometric traits between all three buffalo breeds of

246   the North India outlined the phenotypic distinctness for majority of the body biometric trait. The

247   coefficient of variation (CV) percentage was least for body height in all three breeds. On comparing

248   average of HT, BL, CG, and PG, Gojri buffaloes were found to be of smaller size than Murrah and

249   Nili-Ravi. Nivsarkar et al. (2000) in Nili-Ravi reported average HT, CG, and BL as 134.2, 207.7,

250   and 165.4 cm, respectively, which is comparable to our results. CV% was highest for HL in Gojri

251   (19.52%) and Murrah (12.61%) buffaloes indicating lesser selection pressure on them and more

252   environmental influence. Face width (FW) was least variable in Murrah and Nili-Ravi while it

253   varied greatly in Gojri buffaloes. Most of the body biometric traits measured were less variable

254   indicating their reliability in population classification studies.

255   In the canonical discriminant analysis (refer to table 2), two functions were needed for separation

256   of three distinct population (Asamoah-Boaheng and Sam, 2016) and the first function (function 1)

257   explains 66.7% of the variance and has a Wilk's lambda (0.008) with $p = 0.0001$. The second

258   function explains only 33.3% of the variance in the data, with a recorded $p = 0.0001$ for Wilk's

259   lambda (0.122). Wilks' Lambda value close to zero represents a greater number of variables

260   contribute to the discriminant function (Toalombo Vargas et al., 2019), thus the first function in

261   this study plays major role in classifying the breeds.

262   **Microsatellite variations and Genetic diversity**

263   Microsatellite marker data being the best-suited molecular information for the assessment of

264   genetic diversity (Bowcock et al., 1994; Laval et al., 2000; Groeneveld et al., 2010), allows future

265   management and conservation of the breeds based on their genetic architecture (Luikart et al.,

266   2003; Taberlet, et al., 2008; Toro et al., 2009; Teneva et al., 2013). The FAO and the ISAG/FAO

267   Advisory Group on Animal Genetic Diversity have proposed a panel of 25 SSR markers for

268   diversity studies in buffaloes (Singh et al., 2018). Hence, in the present study the 22 highly

269   polymorphic microsatellite markers out of 25 marker panel, were used for diversity analysis.

270   The mean number of alleles (N_a) in population over a range of loci is considered a fair indicator of

271   allelic variation. The mean N_a ranged from 0-10, 0-13, and 3-14 in Nili-Ravi, Murrah, and Gojri

272   buffaloes, respectively (refer to table 3). The mean N_a per locus for each population in the present

273   study is similar to the reports of Kathiravan et al. (2010) in South Kanara buffaloes; Marques et al.

274   (2011) in Brazilian buffaloes, Martinez et al. (2006), Bhuyan et al. (2010) in Murrah buffaloes and

275   Sajid Ali et al. (2020) in Purnathadi buffaloes. However, a higher number of alleles per locus

276   ranged from 11-26 alleles in Indian water buffaloes have been reported by Vijh et al. (2008). The

277   type of breed under investigation, usage of the particular panel of microsatellite markers, methods

278   of genotyping and the genetic polymorphism within the breed itself greatly influence this variation

279   in the N_a.

280   For microsatellite data, Ohta and Kimura (1973) have established the relationship between the

281   expected homozygosity and its estimator θ, under a pure stepwise mutation model *i.e.* expected

282   homozygosity = $1/\sqrt{1 + 2\theta}$. An estimator of θ can be obtained from microsatellite data by applying

283   the formula, $\theta_H = [1/(1-H_e)^2-1]$ (Excoffier et al., 2010), where H_e is the expected heterozygosity.

284   The mean H_e ranging from 0.14 to 0.81 across all three population over all loci (refer to table 3) is

285   indicative of sufficient polymorphism to measure genetic variation (Takezaki and Nei, 1996). In

286 Gojri buffaloes the expected heterozygosity (H_e) ranged from 0.12 to 0.82 that was comparable

287 with the results reported by Singh et al. (2019). While in both Murrah and Nili-Ravi, it ranged from

288 0 to 0.81 (refer to table 3). Similar high overall mean H_e were reported in Pandharpuri (Khade et

289 al., 2019), Mehsana (Jakhesara et al., 2010), Egyptian (Attia et al., 2014) and Purnathadi (Sajid Ali

290 et al., 2020) buffaloes. The substantially high H_e values implies the presence of high genetic

291 variability in the studied buffalo breeds and suitability of the marker panel for the present study.

292 The average F statistics over 19 loci were $F_{IS} = 0.11744$, $F_{ST} = 0.18252$ and $F_{IT} = 0.27852$ (refer to

293 table 4). In the present study, considerable degree of differentiation has been estimated compared

294 to other buffalo populations from different regions, probably because these populations are

295 genetically distinct. Joshi et al. (2012) reported an F_{ST} value of 7.2% in buffaloes of Indo-gangetic

296 plain, while Vijh et al. (2008) reported a value of 9.69%. However, a comparatively lesser value

297 in eight Indian riverine buffalo was reported by Kumar et al. (2006) which was 3.4%. This value

298 suggested the existence of greater genetic differentiation among North-Indian buffalo breeds than

299 breeds found all over India. A heterozygote deficiency was evident from the positive mean F_{IS}

300 value (0.117>0) indicating low to moderate amount of inbreeding in the population. This could be

301 attributed to assortative mating in small herds owned by farmers, genetic hitchhiking, or the null

302 alleles (Mishra et al., 2008). However, AMOVA over all 22 loci showed 23.59% of variations

303 between populations suggesting the distinctness of all three breeds. The F_{IS} value was found to be

304 4.74%, which is comparable to values obtained in Purnathadi buffaloes (Sajid Ali et al., 2020).

305 The pair-wise F_{ST} values ranged from 0.09 between Murrah and Nili-Ravi to 0.32 between Nili-

306 Ravi and Gojri breeds. The F_{ST} between Murrah and Gojri was 0.25 (Fig. 2a). Least differentiation

307 was found between Nili-Ravi and Murrah (0.09) based on Slatkin linearized F_{ST} while it was

308 highest between Nili-Ravi and Gojri (0.46). Between Murrah and Gojri it was found to be 0.33

309 (Fig. 2b). Nei's distance (d); average within and between populations differentiation is presented

310 in the form of a heat map (Fig. 2c), that shows least distance between Murrah and Nili-Ravi breeds

311 and discriminate Gojri as another population. These results were also in compliance with the results

312 from the factorial correspondence analysis based on molecular data and phylogenetic tree obtained

313 from dissimilarity matrix. In the scatter plot of factorial analysis, Murrah and Nili-Ravi are

314 invariably clustered together. Meanwhile, Gojri was found to be plotted on the opposite side of

315 axis-2 (Fig. 3), yet with more scattering among individuals.

316 The linkage disequilibrium method relies on measures of departure from expected genotype and

317 gametic frequencies, which is the basis for estimation of effective population size (Hill, 1981;

318 Waples, 1991; Luikart et al., 2010). The N_e estimated from microsatellite data reflects the true

319 population distribution of North Indian buffaloes. The comparatively higher N_e of Murrah

320 buffaloes is due to the larger population distribution of the breed in India. The N_e estimates of Gojri

321 population reflects its present status and probable serious inbreeding in future. Hence, the ongoing

322 indiscriminate breeding practices should be shifted to implementation of organized breeding

323 policies focussed on conservation of this distinct breed.

324 **Bayesian genetic structure**

325 Structure software (Pritchard et al., 2000) was used to determine the unbiased structure assuming

326 no prior knowledge regarding the number of breeds. The highest delta K (ΔK) value was calculated

327 as previously described (Evanno et al., 2005). The optimum ΔK value (Fig. 4), was found at $K =$

328 3. For $K=2$, there was no differentiation between Nili-Ravi and Murrah breed. One individual from

329 Murrah population showed significant level of admixture from Gojri population. 99.7% of Gojri

330 buffaloes were classified as a different breed whereas 99.7% and 98.9% of Nili-Ravi and Murrah

331   buffaloes were assigned to one single population, respectively. When K is assumed to be four,

332   Gojri buffaloes are assigned to one distinct cluster with 99% of memberships. Structure assigned

333   all three population into three different breeds when K was assumed to be three (Fig. 5). This

334   indicated that studied populations has well differentiated and possess unique allelic combinations

335   despite being reared in similar geographical regions. However, a low to moderate amount of

336   admixture could be observed in both Murrah and Nili-Ravi population. For K=3, Nili-Ravi showed

337   an average admixture of 3.7% from Murrah and 0.4% from Gojri buffaloes while it was quite high

338   for Murrah with an average admixture of 15.2 and 1.2% from Nili-Ravi and Gojri, respectively.

339   While Gojri population was found to have 99.4% pure blood with an admixture of 0.3% from Nili-

340   Ravi and 0.4% from Murrah. Our results indicate the presence of sufficiently large genetic

341   variability among the North Indian Riverine buffaloes. However, Gojri buffalo populations is

342   unique, compared to Murrah and Nili Ravi buffalos, which were found to be genetically closer

343   than expected. Presently the breeding areas of all these populations are overlapping due to adoption

344   of Murrah as an improver breed for milk production, thus leading to its dominance over Nili-Ravi

345   and Gojri buffalo.

346   **Conclusion**

347   This study demonstrated that the characterization and classification of genetic diversity in Indian

348   buffaloes could be better accomplished through a parallel approach comprising morphometric

349   traits and microsatellite markers. Study of buffalo genetic diversity of Northern India revealed

350   admixture of two major dairy buffalo breeds and a distinct buffalo population was identified. The

351   results obtained provides an opportunity for the design of genetic improvement programs with

352   appropriate choice of breeds for upgrading local non-descript buffaloes along with conservation of

353   unique germplasm. The estimates of effective population size and fixation indices indicate absence

354   of intense systematic selection in past. Further studies involving large populations including

355   samples from other regions of Indian buffalo with FAO recommended microsatellite loci are

356   required to understand the genetic relationships among buffalo genetic resource of India.

357   **Acknowledgement**

358   The authors thank the Directors of ICAR-NDRI and ICAR-NBAGR for their support to carry out

359   the study. Technical support received from Mr. Subhash Chander, T-5, NBAGR, is gratefully

360   appreciated.

361

362   **References:**

363   1. Ali, S. S., Kuralkar, S. V., Das, R., Raina, V., Kataria, R. S. & Vohra, V. (2020).

364   Assessment of genetic diversity and bottleneck in Purnathadi buffaloes using short tandem

365   repeat markers. *Animal Biotechnology*, 1-12.

366   2. Ángel-Marín, P. A., Cardona, H., Moreno-Ochoa, M., & Cerón-Muñoz, M. F. (2010).

367   Analysis of genetic diversity in Colombian buffalo herds. *Revista Colombiana de Ciencias*

368   *Pecuarias* **23**(4), 411-421.

369   3. Asamoah-Boaheng, M. & Sam, E. K. (2016). Morphological characterization of breeds of

370   sheep: a discriminant analysis approach. *SpringerPlus* **5**(1), 69.

371   4. Attia, M., Abou-Bakr, S. & Hafez, Y. M. Genetic polymorphism of seven microsatellite

372   DNA markers in Egyptian buffalo. *Animal Biotechnology* (Cattles, Buffalo) **7** (12), 7.

373   5. Barker, J. S. F. (1999). Conservation of livestock breed diversity. *Animal Genetic*

374   *Resources/Resources génétiques animales/Recursos genéticos animals* **25**, 33-43.

375   6. Barker, J. S. F., Tan, S. G., Selvaraj, O. S., & Mukherjee, T. K. (1997). Genetic variation

376   within and relationships among populations of Asian water buffalo (Bubalus bubalis).

377   *Animal Genetics* **28**(1) 1-13.

378   7. Bhuyan, D.K., Sangwan, M.L., Gole, V.C. & Sethi, R.K. (2010). Studies on DNA

379   fingerprinting in Murrah buffaloes using microsatellite markers.

380   http://nopr.niscair.res.in/handle/123456789/10433

381   8. Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R. & Cavalli-Sforza,

382   L. L. (1994). High resolution of human evolutionary trees with polymorphic

383   microsatellites. *Nature* **368**(6470), 455-457.

384   9. Breno Araújo de Melo, Isabele de Melo Nascimento, Lays Thayse Alves dos Santos,

385   Luciano Gomes de Lima, Filipe Chagas Teodózio de Araújo, Raisa Rodrigues Santos Rios,

386   Alberto de Gusmão Couto & Angelina Bossi Fraga (2018). Body morphometric

387   measurements in Murrah crossbred buffaloes (Bubalus bubalis). *Journal of Applied Animal*

388   *Research* **46**:1. 1307-1312. DOI: 10.1080/09712119.2018.1502669

389   10. Bruford, M. W., Bradley, D. G. & Luikart, G. (2003). DNA markers reveal the complexity

390   of livestock domestication. *Nature Reviews Genetics* **4**(11), 900.

391   11. Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J. & Ovenden, J. R. (2014).

392   NeEstimator v2: re-implementation of software for the estimation of contemporary

393   effective population size (Ne) from genetic data. *Molecular ecology resources* **14**(1), 209-

394   214.

395   12. Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of

396   individuals using the software STRUCTURE: a simulation study. *Molecular ecology* **14**(8),

397   2611-2620.

398   13. Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to

399   perform population genetics analyses under Linux and Windows. *Molecular ecology*

400   *resources* **10**(3), 564-567.

401   14. FAO (2011). Molecular genetic characterization of animal genetic resources. In: FAO

402   Animal Production and Health Guidelines. No. 9, Rome.

403   15. FAO (2012). Phenotypic characterization of animal genetic resources. In: FAO Animal

404   Production and Health Guidelines No. 11, Rome. Available at: http://www.fao.org/3/a-

405   i2686e.pdf

406   16. FAO—Food and Agricultural Organization of the United Nations. FAOSTAT. FAO

407   Statistics Division. (2019). Available online: http://www.fao.org/faostat/en/#data/QA

408   17. Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize

409   population structure. *Molecular ecology resources*, **17**(1), 27-32.

410   18. Gargani, M., Pariset, L., Soysal, M. I., Özkan, E., & Valentini, A. (2010). Genetic variation

411   and relationships among Turkish water buffalo populations. *Animal genetics* **41**(1), 93-96.

412   19. Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R.,

413   Jianlin, H., Finlay, E.K., Groeneveld, E. & Weigend S. (2010). Genetic diversity in

414   livestock breeds. *Animal Genetics* **41**(Suppl 1), 6-31.

415   20. Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R.,

416   Finlay, E.K., Jianlin, H., Groeneveld, E.J.A.G. and Weigend, S., 2010. Genetic diversity in

417   farm animals–a review. *Animal genetics* **41**, 6-31.

418   21. Hall, S. J. & Bradley, D. G. (1995). Conserving livestock breed biodiversity. *Trends in*

419   *ecology & evolution* **10**(7), 267-270.

420   22. Hill, W. G. (1981). Estimation of effective population size from data on linkage

421   disequilibrium. *Genetics Research* **38**(3), 209-216.

422   23. Jakhesara, S. J., Rank, D. N., Kansara, J. D., Parikh, R. C., Vataliya, P. H. & Solanki, J. V.

423   (2010). Microsatellite DNA typing for assessment of genetic variability in the Mehsana

424   buffalo breed of India. *BUFFALO BULLEITN IBIC, KASETSART UNIVERSITY, PO BOX*

425   *1084 BANGKOK 10903, THAILAND* **29**(4), 262.

426   24. Jakobsson, M. & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation

427   program for dealing with label switching and multimodality in analysis of population

428   structure. *Bioinformatics* **23**, 1801–1806.

429   25. Johnson, R. A. & Wichern, D. W. (2002). *Applied multivariate statistical analysis*, Upper

430   Saddle River, NJ: Prentice hall **5**(8).

431   26. Kataria, R. S., Sunder, S., Malik, G., Mukesh, M., Kathiravan, P., & Mishra, B. P. (2009).

432   Genetic diversity and bottleneck analysis of Nagpuri buffalo breed of India based on

433   microsatellite data. *Russian journal of genetics* **45**(7), 826-832.

434   27. Kathiravan, P., Mishra, B. P., Kataria, R. S., Goyal, S., Tripathy, K. & Sadana, D. K.

435   (2010). Short tandem repeat based analysis of genetic variability in Kanarese buffalo of

436   South India. *Russian journal of genetics* **46**(8), 988-993.

437   28. Khade, K. A., Panigrahi, M., Ahmad, S. F., Kumar, P. & Bhushan, B. (2019). Genetic

438   characterization and assessment of diversity in Pandharpuri buffalo breed of India using

439   heterologous microsatellite markers. *Animal biotechnology*, 1-6.

440   29. Kumar, S., Gupta, J., Kumar, N., Dikshit, K., Navani, N., Jain, P. & Nagarajan, M. (2006).

441   Genetic variation and relationships among eight Indian riverine buffalo breeds. *Molecular*

442   *ecology* **15**(3), 593-600.

443   30. Lau, C. H., Drinkwater, R. D., Yusoff, K., Tan, S. G., Hetzel, D. J. S., & Barker, J. S. F.

444   (1998). Genetic diversity of Asian water buffalo (Bubalus bubalis): mitochondrial DNA

445   D-loop and cytochrome b sequence variation. *Animal Genetics* **29**(4) 253-264.

446   31. Laval, G., Iannuccelli, N., Legault, C., Milan, D., Groenen, M. A., Giuffra, E. &

447   Geldermann, H. (2000). Genetic diversity of eleven European pig breeds. *Genetics*

448   *selection evolution* **32**(2), 187.

449   32. Li, M. H., Nogovitsina, E., Ivanova, Z., Erhardt, G., Vilkki, J., Popov, R. & Kantanen, J.

450   (2005). Genetic contribution of indigenous Yakutian cattle to two hybrid populations,

451   revealed by microsatellite variation. *Asian-australasian journal of animal sciences* **18**(5),

452   613-619.

453   33. Luikart, G., England, P. R., Tallmon, D., Jordan, S. & Taberlet, P. (2003). The power and

454   promise of population genomics: from genotyping to genome typing. *Nature reviews*

455   *genetics* **4**(12), 981-994.

456   34. Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K. & Allendorf, F. W. (2010).

457   Estimation of census and effective population sizes: the increasing usefulness of DNA-

458   based approaches. *Conservation Genetics* **11**(2), 355-373.

459   35. Marques, J. R. F., Martínez, A. M., Costa, M. R., Albuquerque, M. S. M., Quiroz, J., Plá,

460   J. L. V. & Bermejo, J. V. D. (2011). Genetic diversity of Brazilian buffaloes (Bubalus

461   bubalis) using DNA microsatellites. *Archivos de zootecnia* **60**(232), 1213-1221.

462   36. Martínez, E., Tirado, J. F., Cerón-Muñoz, M. F., Moreno, M., Montoya, A., Corrales, J. D.

463   & Calvo, S. J. (2009). Genetic characterization of Murrah Buffalo breed in Colombia using

464   microsatellite DNA markers. *Livestock Research for Rural Development* **21**(1).

465   37. Mekkawy, W., Hafez, Y. M., Attia, M., Abdel-Salam, S. A. M., & Abou-Bakr, S. (2012).

466   Association analysis between microsatellite DNA markers and milk yield and its

467   components in Egyptian buffaloes using random regression model. *Egyptian Journal of*

468   *Animal Production* **49**(1), 9-18.

469   38. Metta, M., Kanginakudru, S., Gudiseva, N. & Nagaraju, J. (2004). Genetic characterization

470   of the Indian cattle breeds, Ongole and Deoni (Bos indicus), using microsatellite markers–a

471   preliminary study. *BMC genetics* **5**(1), 16.

472   39. Ministry of Agriculture and Farmers Welfare, GoI. 19th Livestock Census Report. New

473   Delhi, India: Department of Animal Husbandry DAHDF; 2012.

474   40. Mishra, B. P., Kataria, R. S., Bulandi, S. S., Kumar, V. & Mukesh, M. (2008). Genetic

475   diversity in river buffalo (Bubalus bubalis) breeds of central India using heterologous

476   bovine microsatellite markers. *Journal of Applied Animal Research* **33**(2), 159-163.

477   41. Mishra, B.P., Kataria, R.S., Kathiravan, P., Singh, K.P., Sadana, D.K. & Joshi, B.K. (2010).

478   Microsatellite based genetic structuring reveals unique identity of Banni among river

479   buffaloes of Western India. *Livestock Science* **127**(2–3), 257-261

480   42. Nei, M. (1987). *Molecular evolutionary genetics*. Columbia university press, New York,

481   NY, USA.

482   43. Nivsarkar, A.E., Vij, P.K. & Tantia, M.S. (2000). Animal genetic resources of India: Cattle

483   and Buffalo. Directorate of information and publications of Agriculture, ICAR, New Delhi,

484   India.

485   44. Notter, D. R. (1999). The importance of genetic diversity in livestock populations of the

486   future. *Journal of animal science* **77**(1), 61-69.

487   45. Ohta, T. & Kimura, M. (1973). A model of mutation appropriate to estimate the number of

488   electrophoretically detectable alleles in a finite population. *Genetics Research* **22**(2), 201-

489   204.

490   46. Özkan Ünal, E., Soysal, M. İ., Yüncü, E., Dağtaş, N. D., & Togan, İ. (2014). Microsatellite

491   based genetic diversity among the three water buffalo (Bubalus bubalis) populations in

492   Turkey. *Archives Animal Breeding* **57**(1), 1-12.

493   47. Pandey, A. K., Sharma, R., Singh, Y., Prakash, B. & Ahlawat, S. P. S. (2006). Evaluation

494   of genetic variability in Kenkatha cattle by microsatellite markers. *Asian-australasian*

495   *journal of animal sciences* **19**(12), 1685-1690.

496   48. Pandey, A. K., Sharma, R., Singh, Y., Prakash, B. B. & Ahlawat, S. P. S. (2006). Genetic

497   diversity studies of Kherigarh cattle based on microsatellite markers. *Journal of genetics*

498   **85**(2), 117-122.

499   49. Perrier, X., Flori, A. & Bonnot, F. (2003). Data analysis methods In: Hamon P, Seguin M,

500   Perrier X, Glaszmann JC Ed, Genetic diversity of cultivated tropical plants.

501   50. Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure

502   using multilocus genotype data. *Genetics* **155**(2), 945-959.

503   51. Ruane, J. (2000). A framework for prioritizing domestic animal breeds for conservation

504   purposes at the national level: a Norwegian case study. *Conservation Biology* **14**(5), 1385-

505   1393.

506   52. Sambrook, J. & Russel, D.W. (2001). Molecular Cloning, Animal Laboratory manual. Cold

507   Spring Harbor Laboratory press, Cold spring Harbor, New York.

508   53. Sharma, R., Maitra, A., Singh, P. K. & Tantia, M. S. (2013). Genetic diversity and

509   relationship of cattle populations of East India: distinguishing lesser known cattle

510   populations and established breeds based on STR markers. *SpringerPlus* **2**(1), 359.

511   54. Simianer, H. (2005). Decision making in livestock conservation. *Ecological Economics*

512   **53**(4), 559-572.

55. Singh, N. P., Vohra, V., Das, R., Verma, U., Tantia, M. S. & Kataria, R. S. (2019). Elucidating the genetic diversity using SSR based markers in Gojri buffalo. *Indian Journal of Animal Sciences* **89**(5), 522-527.

56. Singh, N. P., Yadav, V., Raina, V., Prakah, R., Pal, S. S. & Baranwal, A. (2018). Heterologous microsatellite markers/SSR used in buffaloes species. *Journal of Pharmacognosy and Phytochemistry* **7**(4), 267-271.

57. Sodhi, M., Mukesh, M., Mishra, B. P., Prakash, B., Ahlawat, S. P. S. & Mitkari, K. R. (2005). Evaluation of genetic differentiation in Bos indicus cattle breeds from Marathwada region of India using microsatellite polymorphism. *Animal biotechnology* **16**(2), 127-137.

58. SPSS. 2001. Statistical Package for Social Sciences. SPSS Inc., 444 Michigan Avenue, Chicago, IL 60611.

59. Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**(2), 437-460.

60. Tajima, F. (1993). Measurement of DNA polymorphism. *Mechanisms of Molecular Evolution, In: Introduction to Molecular Paleopopulation Biology*, 37-59.

61. Takezaki, N. & Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, **144**(1), 389-399.

62. Tantia, M. & Vijh, R. (2012). Microsatellite analysis of buffaloes of Indo-Gangetic Plains. *Indian Journal of Animal Sciences* **82**(11), 1434-1437.

63. Teneva, A., Dimitrov, K., Caro Petrović, V., Petrović, M. P., Dimitrova, I., Tyufekchiev, N. & Petrov, N. (2013). Molecular genetics and SSR markers as a new practice in farm animal genomic analysis for breeding and control of disease disorders. *Biotechnology in Animal Husbandry* **29**(3), 405-429.

536   64. Toalombo Vargas, P. A., León, J. M., Fiallos Ortega, L. R., Martinez, A., Villafuerte

537   Gavilanes, A. A., Delgado, J. V. & Landi, V. (2019). Deciphering the Patterns of Genetic

538   Admixture and Diversity in the Ecuadorian Creole Chicken. *Animals* **9**(9), 670.

539   65. Toro, M. & Caballero, A. (2005). Characterization and conservation of genetic diversity in

540   subdivided populations. Philosophical Transactions of the Royal Society of London, Series

541   B: Biological Sciences **360**, 1367–78.

542   66. Toro, M. A., Fernández, J. & Caballero, A. (2009). Molecular characterization of breeds

543   and its use in conservation. *Livestock Science* **120**(3), 174-195.

544   67. Vijh, R. K., Tantia, M. S., Mishra, B. & Bharani Kumar, S. T. (2008). Genetic relationship

545   and diversity analysis of Indian water buffalo (Bubalus bubalis). *Journal of animal science*

546   **86**(7), 1495-1502.

547   68. Vohra, V., Niranjan, S. K. & Joshi, B. K. (2012). Gojri–A novel migratory buffalo

548   germplasm in Punjab and Himachal Pradesh. *Journal of Animal Research* **2**(3), 317-321.

549   69. Vohra, V., Niranjan, S. K., Mishra, A. K., Jamuna, V., Chopra, A., Sharma, N. & Jeong,

550   D. K. (2015). Phenotypic characterization and multivariate analysis to explain body

551   conformation in lesser known buffalo (Bubalus bubalis) from North India. *Asian-*

552   *Australasian journal of animal sciences* **28**(3), 311.

553   70. Vohra, V., Sodhi, M., Niranjan, S. K., Mishra, A. K., Chopra, A., Kumar, M. and Joshi A

554   K. (2017) Characterization of rare migratory cattle and evaluation of its phylogeny using

555   short-tandem-repeat-based markers. *Journal of Applied Animal Research* **45**(1): 355-363

556   71. Waples, R. S. (1991). Genetic methods for estimating the effective size of cetacean

557   populations. *Report of the International Whaling Commission (special issue)* **13**, 279-300.

558   72. Weitzman, M. L. (1993). What to preserve? An application of diversity theory to crane

559   conservation. *The Quarterly Journal of Economics* **108**(1), 157-183.

560   73. Yadav, D. K., Arora, R. & Jain, A. (2017). Classification and conservation priority of five

561   Deccani sheep ecotypes of Maharashtra, India. *PloS one* **12**(9).

562   74. Yoon, D. H., Kong, H. S., Oh, J. D., Lee, J. H., Cho, B. W., Kim, J. D. & Lee, H. K. (2005).

563   Establishment of an individual identification system based on microsatellite

564   polymorphisms in Korean cattle (Hanwoo). *Asian-australasian journal of animal*

565   *sciences* **18**(6), 762-766.

566   75. Zhang, Y., Sun, D., Yu, Y., & Zhang, Y. (2007). Genetic diversity and differentiation of

567   Chinese domestic buffalo based on 30 microsatellite markers. *Animal Genetics* **38**(6), 569-

568   575.

569   76. Zhao, G., & Maclean, A. L. (2000). A comparison of canonical discriminant analysis and

570   principal component analysis for spectral transformation. *PE&RS, Photogrammetric*

571   *Engineering & Remote Sensing* **66**(7), 841-847.

572   77. Zouros, E. (1979). Mutation rates, population sizes and amounts of electrophoretic

573   variation of enzyme loci in natural populations. *Genetics* **92**(2), 623-646.

575

Average measurements of body morphometric traits in 3 buffalo populations from Northern India

Pop, Mu, NR, Goj in the table corresponds to 'Population', 'Murrah', 'Nili-Ravi' and 'Gojri', respectively. * The mean difference is significant at the 0.05 level.

Traits (measured in cm)	Pop	Mean ± SE	SD	Min.	Max.	CV%	Pop (i)	Pop (j)	Mean Difference (i-j)	*p*-value
Body Height	Mu	138.40±0.41	4.40	129.00	150.00	3.17	Mu	NR	4.29±0.85	0.095
	NR	134.10±1.10	4.79	108.00	134.00	4.35	Mu	Goj	9.58*±0.63	0.0001
	Goj	128.82±0.47	4.51	118.00	145.00	3.49	NR	Goj	-5.28*±0.88	0.0001
Body Length	Mu	129.26±0.55	5.85	115.00	147.00	4.52	Mu	NR	23.45*±1.08	0.0001
	NR	105.81±1.12	6.83	91.00	121.00	6.45	Mu	Goj	-4.22*±0.80	0.0001
	Goj	133.48±0.51	4.90	122.00	151.00	3.67	NR	Goj	-27.67*±1.11	0.0001
Chest Girth	Mu	212.53±1.12	11.87	185.00	250.00	5.58	Mu	NR	51.37*±2.01	0.0001
	NR	161.16±1.52	9.27	144.00	182.00	5.74	Mu	Goj	16.63*±1.49	0.0001
	Goj	195.90±0.98	9.43	170.00	214.00	4.81	NR	Goj	-34.74*±2.07	0.0001
Paunch Girth	Mu	232.11±1.09	11.54	208.00	266.00	4.97	Mu	NR	60.84*±2.84	0.0001
	NR	171.27±1.79	10.89	153.00	198.00	6.35	Mu	Goj	18.91*±2.11	0.0001
	Goj	213.20±2.03	19.50	121.00	242.00	9.14	NR	Goj	-41.92*±2.92	0.0001
Face Length	Mu	49.29±0.25	2.61	46.00	62.00	5.27	Mu	NR	8.40*±0.41	0.0001
	NR	40.89±0.31	1.88	38.00	46.00	4.60	Mu	Goj	0.66±0.31	0.081
	Goj	48.63±0.17	1.66	44.00	54.00	3.41	NR	Goj	-7.74*±0.43	0.0001
Face Width	Mu	19.72±0.09	0.98	18.00	22.00	4.92	Mu	NR	-0.15±0.51	0.955
	NR	19.86±0.19	1.18	18.00	23.00	5.95	Mu	Goj	-3.29*±0.38	0.0001
	Goj	23.01±0.43	4.17	20.00	49.00	18.10	NR	Goj	-3.15*±0.53	0.0001
Ear Length	Mu	28.41±0.11	1.22	25.00	30.00	4.29	Mu	NR	7.92*±0.23	0.0001
	NR	20.49±0.17	1.04	19.00	22.00	5.09	Mu	Goj	-0.30±0.17	0.110
	Goj	28.75±0.13	1.25	21.00	31.00	4.34	NR	Goj	-8.26*±0.24	0.0001
Horn Length	Mu	28.37±0.34	3.59	16.00	34.00	12.61	Mu	NR	-17.84*±1.19	0.0001
	NR	46.22±0.88	5.32	34.00	56.00	11.51	Mu	Goj	-16.36*±0.88	0.0001
	Goj	44.73±0.91	8.73	23.00	82.00	19.52	NR	Goj	1.49±1.22	0.443
Horn Circumference	Mu	17.18±0.16	1.70	12.00	21.00	9.90	Mu	NR	-2.31*±0.33	0.0001
	NR	19.49±0.29	1.76	17.00	25.00	9.02	Mu	Goj	-2.69*±0.24	0.0001
	Goj	19.87±0.18	1.73	17.00	28.00	8.70	NR	Goj	-0.38±0.34	0.488
Hip Bone	Mu	55.55±0.29	3.13	49.00	63.00	5.61	Mu	NR	15.95*±0.6	0.0001
	NR	39.59±0.36	2.20	34.00	43.00	5.56	Mu	Goj	1.95*±0.45	0.0001
	Goj	53.60±0.37	3.53	30.00	60.00	6.69	NR	Goj	-14.00*±0.62	0.0001
Pin Bone	Mu	16.94±0.10	1.08	15.00	20.00	6.32	Mu	NR	2.83*±0.54	0.0001
	NR	14.11±0.30	1.81	10.00	17.00	12.80	Mu	Goj	-7.37*±0.40	0.0001
	Goj	24.30±0.45	4.31	19.00	59.00	17.69	NR	Goj	-10.20*±0.55	0.0001
Tail Length	Mu	104.61±1.24	13.18	76.00	130.00	12.58	Mu	NR	24.26*±2.59	0.0001
	NR	80.35±1.06	6.42	69.00	102.00	7.99	Mu	Goj	13.65*±1.92	0.0001
	Goj	90.96±1.68	16.11	22.00	116.00	17.70	NR	Goj	-10.60*±2.66	0.0001
Tail up to Switch	Mu	95.14±1.04	11.01	68.00	123.00	11.56	Mu	NR	25.87*±1.84	0.0001
	NR	69.27±0.98	5.93	60.00	85.00	8.56	Mu	Goj	-9.14*±1.37	0.0001
	Goj	104.28±0.97	9.27	73.00	124.00	8.87	NR	Goj	-35.01*±1.9	0.0001

Characteristics of canonical discriminant functions and test statistics

[†] Degrees of freedom

Discriminant Function	Eigen values	Variance percentage explained	Cumulative variance	Canonical correlation	Wilks' Lambda	Chi-square	†d.f.	P
1st function	14.40	66.7	66.7	0.967	0.008	1134.64	20	0.0001
2nd function	7.19	33.3	100	0.937	0.122	493.30	9	0.0001

Breed wise details of estimated Genetic Diversity Indices for each microsatellite markers

NE: not estimable in the global analysis of molecular variance

Locus	Expected Heterozygosity (H_e)			Theta H (θ_H)			Number of alleles (N_a)			F_{ST}	p-value of F_{ST} estimation
	Nilli-Ravi	Murrah	Gojri	Nilli-Ravi	Murrah	Gojri	Nilli-Ravi	Murrah	Gojri		
BM1818	0.69	0.68	0.71	2.29	2.17	2.49	7	9	6	0.029	0.078
CSSM19	0.74	0	0.73	2.80	0	2.75	6	0	6	NE	NE
CSSM33	0.71	0.73	0.63	2.49	2.67	1.74	8	9	7	0.194	0.00001
CSSM45	0.65	0.80	0.73	1.87	3.98	2.72	5	6	6	0.204	0.00001
CSSM47	0.81	0.69	0.82	4.23	2.25	4.63	10	10	11	0.175	0.00001
CSSM66	0.79	0.61	0.82	3.67	1.60	4.43	7	6	9	0.176	0.00001
Hel013	0.67	0.75	0.82	2.08	3	4.49	8	8	9	0.139	0.00001
ILSTS19	0.14	0.17	0.12	0.16	0.20	0.14	2	6	3	0.809	0.00001
ILSTS25	0.58	0.61	0.72	1.39	1.59	2.52	5	6	8	0.179	0.00001
ILSTS26	0.67	0.61	0.76	2.05	1.57	3.26	6	6	5	0.084	0.00001
ILSTS28	0.76	0.76	0.76	3.25	3.16	3.21	8	7	6	0.002	0.7165
ILSTS29	0.33	0.26	0.82	0.49	0.35	4.49	6	4	10	0.376	0.00001
ILSTS30	0.71	0.60	0.69	2.45	1.47	2.25	7	6	6	0.194	0.00001
ILSTS33	0.66	0.68	0.59	1.99	2.14	1.47	6	7	3	0.005	0.51026
ILSTS36	0.66	0.67	0.72	1.98	2.03	2.58	6	6	12	0.160	0.00001
ILSTS52	0.67	0	0.73	2.00	0	2.72	9	0	8	NE	NE
ILSTS56	0.40	0.53	0.59	0.67	1.12	1.45	5	7	8	0.345	0.00001
ILSTS58	0.81	0.85	0.77	4.24	5.70	3.40	7	9	11	0.088	0.00001
ILSTS60	0.45	0.36	0.60	0.81	0.57	1.53	4	3	14	0.016	0.18573
ILSTS61	0.66	0.81	0.76	1.91	4.23	3.23	6	13	11	0.178	0.00001
ILSTS089	0	0.79	0.76	0	3.76	3.14	0	6	6	NE	NE
ILSTS95	0.80	0.71	0.71	4.08	2.43	2.48	10	11	8	0.061	0.00001
Mean	0.61	0.58	0.70	1.55	1.36	2.33	6.27	6.59	7.86	0.18	-
SD	0.22	0.25	0.15	-	-	-	2.33	3.10	2.83	-	-

Results of Global Molecular Analysis of Variance (AMOVA) along with fixation indices in Northern India buffalo populations

Sources of Variation	Degrees of Freedom	Sum of Squares	Variance components	Variation explained (%)	Fixation indices	P Value
Among populations	2	232.21	1.38	18.25	0.182 (F_{ST})	0.00001
Among individuals within populations	125	786.65	0.73	9.60	0.117 (F_{IS})	0.00001
Within individuals	128	640.50	5.46	72.15	0.278 (F_{IT})	0.00001
Total	255	1659.37	7.57	100	-	-

Individual wise ancestry level inferred through Bayesian method in Structure

NR, MU, and Goj represent Nili-Ravi, Murrah, and Gojri, respectively

a

Label	Population	Inferred clusters		
		Goj	NR	MU
NR1	1	0.002	0.969	0.029
NR2	1	0.002	0.988	0.011
NR3	1	0.008	0.958	0.034
NR4	1	0.004	0.958	0.038
NR5	1	0.001	0.99	0.009
NR6	1	0.002	0.994	0.004
NR7	1	0.003	0.991	0.005
NR8	1	0.002	0.992	0.006
NR9	1	0.004	0.975	0.021
NR10	1	0.002	0.994	0.004
NR11	1	0.002	0.987	0.011
NR12	1	0.002	0.992	0.006
NR13	1	0.005	0.987	0.007
NR14	1	0.004	0.991	0.005
NR15	1	0.001	0.928	0.071
NR16	1	0.002	0.601	0.397
NR17	1	0.001	0.993	0.006
NR18	1	0.004	0.959	0.037
NR19	1	0.002	0.993	0.005
NR20	1	0.002	0.947	0.051
NR21	1	0.003	0.984	0.013
NR22	1	0.005	0.99	0.005
NR23	1	0.004	0.99	0.006
NR24	1	0.002	0.993	0.005
NR25	1	0.002	0.991	0.007
NR26	1	0.004	0.993	0.003
NR27	1	0.020	0.973	0.008
NR28	1	0.002	0.996	0.003
NR29	1	0.001	0.995	0.004
NR30	1	0.003	0.981	0.015
NR31	1	0.002	0.541	0.457
NR32	1	0.002	0.993	0.005
NR33	1	0.002	0.958	0.04
NR34	1	0.002	0.993	0.006
NR35	1	0.001	0.993	0.006
NR36	1	0.003	0.994	0.004
NR37	1	0.001	0.988	0.01
NR38	1	0.003	0.986	0.011
NR39	1	0.04	0.843	0.118

NR40	1	0.003	0.994	0.003
MU1	2	0.004	0.015	0.981
MU2	2	0.28	0.707	0.013
MU3	2	0.003	0.003	0.993
MU4	2	0.005	0.007	0.988
MU5	2	0.001	0.003	0.995
MU6	2	0.005	0.013	0.983
MU7	2	0.008	0.048	0.944
MU8	2	0.004	0.946	0.05
MU9	2	0.002	0.003	0.995
MU10	2	0.003	0.01	0.987
MU11	2	0.002	0.004	0.994
MU12	2	0.002	0.119	0.879
MU13	2	0.003	0.229	0.768
MU14	2	0.003	0.461	0.536
MU15	2	0.005	0.004	0.99
MU16	2	0.005	0.019	0.976
MU17	2	0.01	0.303	0.687
MU18	2	0.002	0.018	0.979
MU19	2	0.039	0.36	0.602
MU20	2	0.002	0.006	0.992
MU21	2	0.003	0.003	0.995
MU22	2	0.004	0.804	0.192
MU23	2	0.007	0.164	0.829
MU24	2	0.003	0.007	0.99
MU25	2	0.003	0.004	0.993
MU26	2	0.002	0.065	0.933
MU27	2	0.001	0.011	0.988
MU28	2	0.017	0.084	0.899
MU29	2	0.022	0.049	0.929
MU30	2	0.002	0.011	0.987
MU31	2	0.002	0.007	0.991
MU32	2	0.014	0.438	0.548
MU33	2	0.002	0.004	0.995
MU34	2	0.002	0.009	0.99
MU35	2	0.004	0.02	0.975
MU36	2	0.004	0.983	0.013
MU37	2	0.002	0.004	0.994
MU38	2	0.005	0.015	0.98
MU39	2	0.005	0.104	0.891
MU40	2	0.002	0.01	0.988

Goj1	3	0.995	0.002	0.003
Goj2	3	0.995	0.002	0.003
Goj3	3	0.996	0.002	0.001
Goj4	3	0.991	0.006	0.003
Goj5	3	0.996	0.002	0.002
Goj6	3	0.993	0.004	0.002
Goj7	3	0.997	0.001	0.002
Goj8	3	0.988	0.008	0.004
Goj9	3	0.995	0.002	0.003
Goj10	3	0.989	0.006	0.005
Goj11	3	0.997	0.001	0.002
Goj12	3	0.997	0.001	0.002
Goj13	3	0.997	0.001	0.002
Goj14	3	0.997	0.001	0.002
Goj15	3	0.991	0.003	0.006
Goj16	3	0.995	0.002	0.002
Goj17	3	0.996	0.002	0.002
Goj18	3	0.995	0.003	0.002
Goj19	3	0.991	0.005	0.004
Goj20	3	0.986	0.009	0.005
Goj21	3	0.997	0.001	0.002
Goj22	3	0.996	0.002	0.002
Goj23	3	0.99	0.004	0.006
Goj24	3	0.997	0.001	0.001
Goj25	3	0.996	0.002	0.002
Goj26	3	0.996	0.002	0.002
Goj27	3	0.997	0.002	0.002
Goj28	3	0.995	0.002	0.002
Goj29	3	0.995	0.002	0.003
Goj30	3	0.995	0.002	0.002
Goj31	3	0.995	0.002	0.003
Goj32	3	0.994	0.003	0.004
Goj33	3	0.992	0.004	0.004
Goj34	3	0.996	0.002	0.002
Goj35	3	0.995	0.002	0.003
Goj36	3	0.995	0.002	0.003
Goj37	3	0.992	0.004	0.005
Goj38	3	0.995	0.002	0.003
Goj39	3	0.992	0.001	0.006
Goj40	3	0.996	0.002	0.002
Goj41	3	0.994	0.003	0.004

Goj42	3	0.967	0.006	0.026
Goj43	3	0.989	0.003	0.008
Goj44	3	0.994	0.003	0.003
Goj45	3	0.988	0.006	0.006
Goj46	3	0.997	0.001	0.002
Goj47	3	0.993	0.003	0.004
Goj48	3	0.991	0.002	0.007

Figure 1

Canonical Discriminant Analysis (Scatter Plot) based on 13 body morphometric traits depicted three different buffalo populations from Northern India
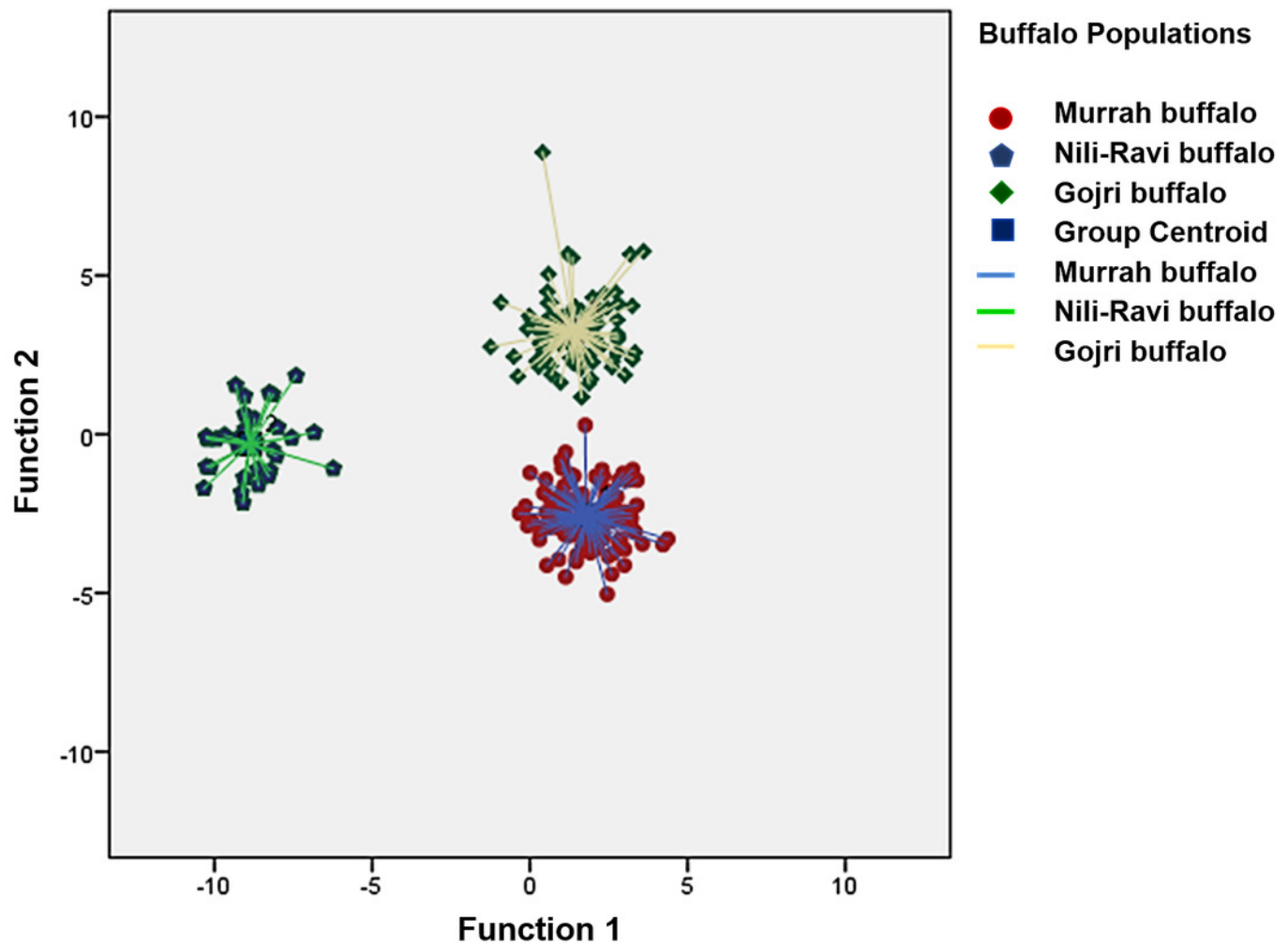
Figure 2

Heatmap of molecular diversity indices

(a) pairwise $F_{ST,}$ (b) Slatkin's linearized $F_{ST,}$ (c) Nei's distance and AMOVA
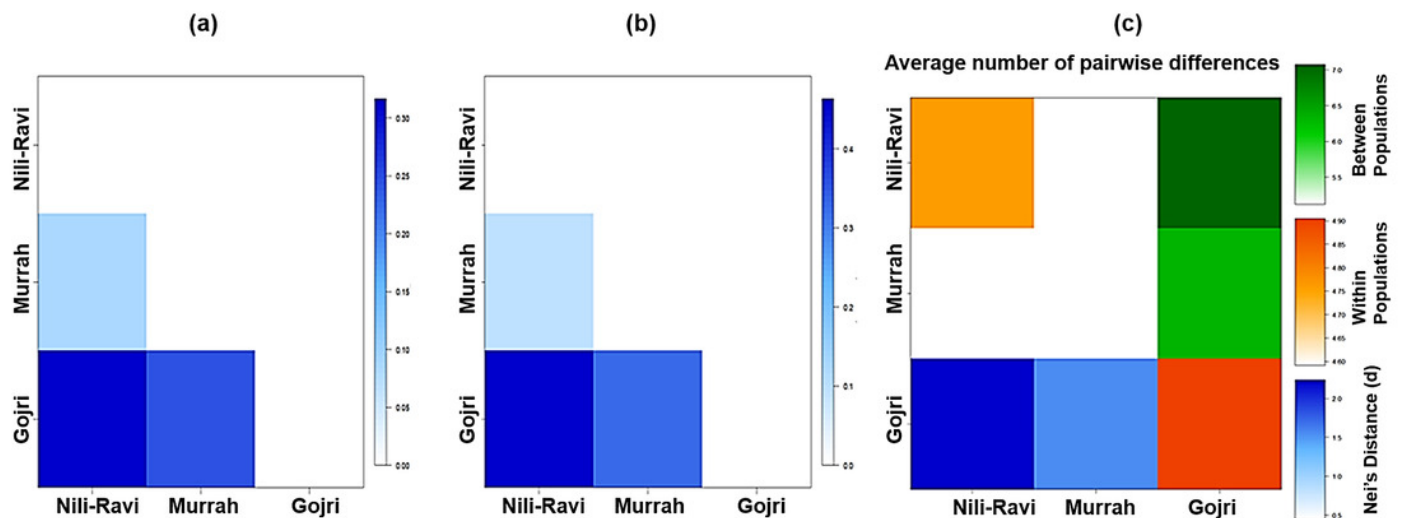
Figure 3

Scatter plot for Factorial Correspondence Analysis based on genetic diversity indices
depicted three different buffalo populations from Northern India
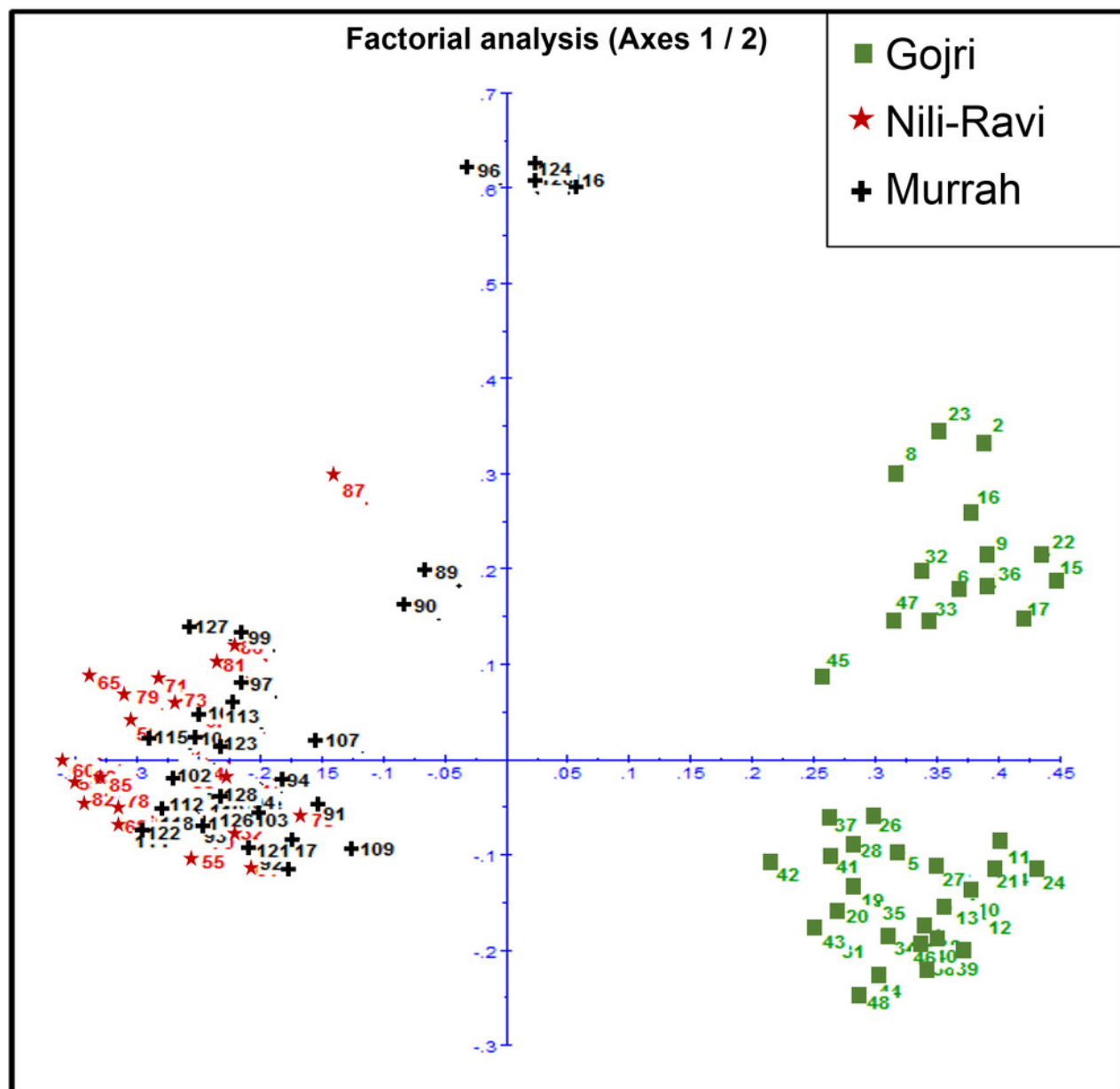
Figure 4

Estimates of number of sub-populations (*K*) using different statistics by Evanno method to determine ideal number of clusters present in the studied buffalo populations
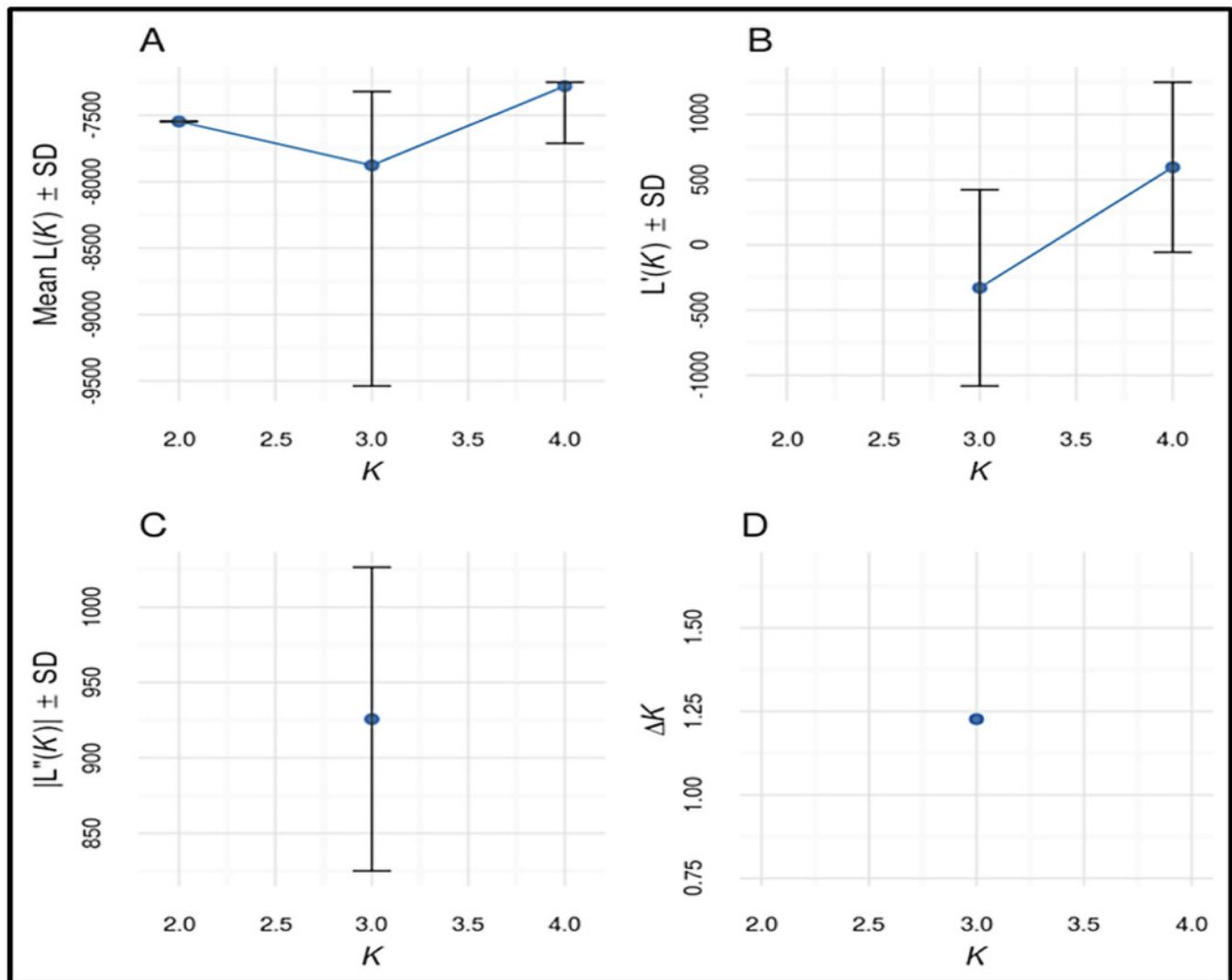
Figure 5

Bayesian clustering of North-Indian buffalo populations under the assumption of K = 2–4 using STRUCTURE program reveals genetic admixture and introgression among Murrah and Nili-Ravi populations while Gojri buffalo is genetically distinct.

Each vertical bar represents individuals displaying membership coefficients for each population cluster. Populations are separated by dashed white lines. Graphics were obtained with CLUMPP (Jakobsson & Rosenberg, 2007).