Morphometric and microsatellite-based comparative genetic diversity analysis in *Bubalus bubalis* from North India (#59778)

First revision

Guidance from your Editor

Please submit by **2 Jul 2021** for the benefit of the authors .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
8 Figure file(s)
5 Table file(s)
1 Raw data file(s)

Custom checks

Vertebrate animal usage checks

- Have you checked the authors ethical approval statement?
- Were the experiments necessary and ethical?
- Have you checked our animal research policies?

Field study

- Have you checked the authors field study permits?
- Are the field study permits appropriate?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

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

A parallel approach of morphometric characterization and molecular diversity has been used to classify the buffalo germplasm of Northern India. Diversity analysis of the morphometric information revealed different clusters suggesting distinct genetic entities among the studied populations. Molecular diversity was analyzed, using a panel of 22 microsatellite DNA markers. Analysis of molecular variance revealed 81.8% of genetic variance was found within breeds, while 18.8% of the genetic variation was found among breeds that could differentiate the studied buffalo breeds into 3 sub-populations. Effective population sizes for each breed estimated based on linkage disequilibrium were 142, 75, and 556 in Gojri, Nili-Ravi, and Murrah populations, respectively. 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. Molecular diversity analysis suggested admixture of the Murrah and Nili-Ravi while labelled the Gojri as a unique population. The study provides important information on the North-Indian buffaloes that could be utilized in designing their breeding, improvement and conservation programs.

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

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

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

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

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

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

[†]**Corresponding author:**

Dr. Vikas Vohra, Principal Scientist

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

Email: Vikas.Vohra@icar.gov.in

vohravikas@gmail.com

Abstract

A parallel approach of morphometric characterization and molecular diversity has been used to classify the buffalo germplasm of Northern India. Diversity analysis of the morphometric information revealed different clusters suggesting distinct genetic entities among the studied populations. Molecular diversity was analyzed, using a panel of 22 microsatellite DNA markers. Analysis of molecular variance revealed 81.8% of genetic variance was found within breeds, while 18.8% of the genetic variation was found among breeds that could differentiate the studied buffalo breeds into 3 sub-populations. Effective population sizes for each breed estimated based on linkage disequilibrium were 142, 75, and 556 in Gojri, Nili-Ravi, and Murrah populations, respectively. 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. Molecular diversity analysis suggested admixture of the Murrah and Nili-Ravi while labelled the Gojri as a unique population. The study provides important information on the North-Indian buffaloes that could be utilized in designing their breeding, improvement and conservation programs.

Keywords: Buffalo, Genetic Variation, India, Microsatellite Markers

Introduction

India is the largest milk-producing country in the world and buffaloes alone contributes around half (49%) to the milk production of India. Water buffalo *(Bubalus Bubalis*) that most probably domesticated in Indus Valley region for multiple utility creates a rich Bubaline diversity in Northern regions of India. Buffaloes are producing around half (49%) of the total milk produced by world's top milk producer country i.e., India. North India, comprising of Punjab, Haryana, Himachal Pradesh, Delhi and Western Uttar Pradesh, is the largest milk-producing region in the country. The predominant bubaline genetic resources documented from the region include Murrah, Nili Ravi and Gojri buffaloes (http://www.nbagr.res.in/regbuf.html). Murrah being dominating buffalo germplasm with superior milk-producing ability has suppressed the need for identification and characterization of other breeds. On the other hand, Gojri is one of the little-known buffalo population of the region, with a good milch potential on low to zero input system of dairying and is maintained on a semi-migratory extensive system of management (Vohra et al., 2012 and 2015).

Characterization and classification of animal genetic resources (AnGR) require ample knowledge of the geographical distribution of the breeds, identification of unique characteristics, population size and structure, production environment, and genetic diversity. It is customary to perform a detailed molecular study along with physical and phenotypic assessment to check within and between population diversity in order to characterize a population (Weitzman, 1993; Hall & Bradley, 1995; Barker, 1999; Ruane, 2000; Bruford et al., 2003; Simianer, 2005; Toro & Caballero, 2005). Vohra et al. (2015) have used 13 morphometric traits of Gojri buffaloes for phenotypic characterization using Principal Component Analysis, a multivariate statistical technique. Multivariate statistical analysis techniques viz. classical principal component analysis

serves the objectives of dimension reduction and clustering when multiple morphometric traits are measured (Johanson and Wichern, 2002; Yadav et al., 2017).

The neutrality, co-dominant inheritance and high polymorphic information content of microsatellite markers have rendered them as the markers of choice for diversity studies (Metta et al., 2004; Li et al., 2005; Yoon et al., 2005; Sodhi et al., 2005, 2006; Kumar et al., 2006; Pandey et al., 2006a; Pandey et al., 2006b; Vijh et al., 2008; Sharma et al., 2013). The genetic diversity within Murrah, Nili-Ravi and Gojri breeds have been studied independently that share the common breeding tract in North India. However, in India, buffalo breeding is largely restricted to natural mating that subsequently may have led to admixture of these populations. Hence, there is a need to assess the between-breed genetic diversity among these breeds. The present study was performed to assess the levels of genetic diversity, and population structure among three buffalo breeds of North India. The results will help in formulating an effective breeding, management policy, shaping future conservation plans for maintaining breed purity and reducing the possible admixture due to introgression among purebreds. Thus, it is imperative to compare the region-specific diversity and breed status of bubaline germplasm.

Materials & Methods

Sampling strategy

Sampling was done from their respective native tracts, to compare the genetic diversity between three different breeds. Gojri buffalo samples were collected during 2017-18 from areas of Punjab 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-Ravi buffaloes were collected from Punjab state (28° 17' to 32° 17' N and 74° to 76° 41' E). The Nili-Ravi has a comparatively smaller geographical distribution compared to Murrah and Gojri. In

India, Murrah buffaloes are found in almost all regions but its native area is Haryana state (28^0 02′ to 30^0 21′ N and 75^0 to 77^0 E) hence, sampling was performed from Haryana and Punjab. The data of Murrah and Nili-Ravi was taken for comparative analysis from Buffalo Genomics Lab of National Bureau of Animal Genetic Resources, Karnal. The breeding and sampling tract had a herd size of 2-6 buffaloes per households. To ensure that selected animals are unrelated, in the absence of detailed pedigree accounts, buffalo breeders were interviewed in detail and their records were checked. Only those animals who were not having common parents for at least 3-4 generations were included in the study. Buffaloes were selected for this study following guidelines of measurement of domestic animal diversity program (FAO, 2011) those represented the original indigenous true to type phenotype. Blood samples were collected with the consent of herd owners. Approximately 5-10 ml of blood from jugular vein was collected by trained Veterinarian using aseptic measures. All the studies were carried out under approval of ICAR-National Dairy Research Institute IAEC 1705/GO/ac/13/CPCSEA.

Morphometric traits were measured on a total of 242 adult female buffaloes, comprising of 113 Murrah, 37 Nili-Ravi, and 92 Gojri buffaloes, to avoid the sex and age differences. Thirteen (13) different traits were measured on all three breeds as suggested by Breno et al (2018). All the measurements on the animal were recorded in their normal standing position on a levelled surface using a tape measure by the same technical person. Traits recorded were body height (HT), body length (BL), chest girth (CG), paunch girth (PG), face length (FL), face width (FW), horn length (HL), horn circumference (HC), ear length (EL), distance between hip bone (HB), distance between pin bone (PB), tail length (TL), and tail length up to switch (TS). To avoid age effects, only adult buffaloes (3.5 years above) were included in study. For microsatellite genotyping, blood samples were collected from 128 (40 Murrah, 40 Nili-Ravi, and 48 Gojri) buffaloes.

Genotyping microsatellite markers

Genomic DNA was isolated from blood samples by standard phenol–chloroform extraction protocol, as described by Sambrook and Russel (2001). DNA concentration was checked by spectrophotometric method. Genetic variation was assayed using 25 microsatellite markers. Microsatellite genotyping was carried out as previously describe in Vohra et al. (2017) following the protocol of Mishra et al. (2010). Fluorescent-tagged forward primers for each microsatellite were used. The primers those were able to produce a fragment size >75bp were used in the study. Fragment length analysis was performed through ABI PRISM 3100 automatic sequencer (Applied Biosystems, Foster City, CA, USA) after performing polymerase chain reaction (PCR) for fragment amplification. Allele length for the different fragments generated was determined as described in Vohra et al. (2017) using GeneScan software (version 5.0 Applied Bio system). Observed number of alleles (N_a), theta estimate (θ_H), expected heterozygosity (H_e), F_{IT} (total inbreeding estimate), F_{ST} (measurement of population differentiation) and F_{IS} (within- population-inbreeding estimate) were calculated using Arlequin v3.5 (Excoffier et al., 2010). Pairwise differences between populations using molecular distances were calculated. Molecular diversity indices were calculated as per Tajima (1983), Tajima (1993), Nei (1987), and Zouros (1979), implemented in Arlequin v3.5, and allowing 5% level of missing data. Analysis of molecular variances was done using 1,000 permutations. Exact test of population differentiation was performed with 1,00,000 Markov chain steps and 10,000 dememorization steps.

Statistical analysis

Statistical analyses on morphometric data were performed using SPSS v17.0 software (SPSS, 2001). The canonical discriminant analysis was performed in SAS v9.3 program (SAS Institute

Inc., 2011) using Proc disc procedure, for determining the most discriminatory morphometric traits. The probabilities of assigning an individual to a population were determined using Discrim procedure based on the linear discriminant function that included the thirteen morphometric variables. Wilk's Lambda was used as the test statistics to check for the differences between the means of identified groups of subjects on a combination of dependent variables.

Population assignment was performed using the Bayesian Markov chain Monte Carlo approach implemented in Structure v2.3.4 (Pritchard et al., 2000). The most likely number of subpopulations was determined by the Evanno ΔK method (Evanno et al., 2005) using R package "POPHELPER" (Francis et al., 2017). Twenty independent runs were performed for $K = 2$ to 4 to identify the most likely number of clusters present in the dataset. The analysis was performed with a burn in period of 10000 and 50000 MCMC iterations. Effective population size (N_e) was checked for the three population. N_e was estimated using linkage disequilibrium method using NeEstimator v2.01 (Do et al., 2014) Software. The P-critical value (rare allele frequency) was set to 0.05, below which all the alleles were rejected. Jackknife confidence intervals (CI) were calculated for each estimate, N_e, of different population. Discrimination between populations was elucidated graphically through principal coordinate analysis (PCoA) using Darwin v6.0.021 (Perrier et al., 2003). The dissimilarity matrix based phylogenetic tree was also obtained through Darwin.

Results

Classificatory analysis based on Morphometric traits

The means and standard deviation, coefficient of variations and comparison of mean difference between populations for each trait across population is listed in Table 1. A Canonical Discriminant analysis was used to compare different morphometric traits and first two canonical discriminant

functions were used in the analysis, which explained 66.7% and 33.3% of total variance, respectively. Wilk's Lambda was used as the test statistics to check the difference between means of the two groups and was found to be significant (Table 2). Classification based on canonical discriminant functions for both original and cross-validated counts predicted 100% assignment of each adult buffaloes to their hypothetically known populations *i.e.* Murrah, Nili-Ravi, and Gojri. All the individuals plotted based on 1st and 2nd canonical discriminant functions were clustered into three distinct groups suggesting three different breeds in the sample (Fig. 1).

Microsatellite variations

Among 25 microsatellite loci genotyped for this study, only 22 loci that were polymorphic for all three populations were used for further downstream analysis. A total of 145, 138, and 173 alleles were found across 22 loci in the 128 individuals sampled from the Murrah, Nili-Ravi, and Gojri buffaloes, respectively. ILSTS60 was highly polymorphic in Gojri buffaloes, ILSTS95 in both Murrah and Nili-Ravi and ILSTS61 in Murrah (Fig. S1a). Mean number of alleles for all populations varied from 3.67±2.08 at ILSTS19 to 10.33±0.58 at CSSM47. Mean expected heterozygosity (H_e) across all populations ranged from 0.14±0.02 at ILSTS19 to 0.81±0.04 at ILSTS58. The mean H_e estimated over all loci was lowest in Murrah (0.58±0.25) while it was highest in Gojri population (0.70±0.15) (Fig. S1b). Estimator of mutation parameter (θ_H) that is obtained using observed homozygosity values was estimated under infinite allele model. Mean θ_H ranged from 1.36 in Murrah to 2.33 in Gojri buffaloes (Fig. S1c). Across all three populations mean θ_H ranged from 0.17±0.03 (ILSTS19) to 4.44±1.16 (ILSTS58). Marker wise number of alleles, H_e, and θ_H in each breed given in Table 3.

Genetic diversity

Global Analysis of molecular variance (AMOVA) using 19 polymorphic loci was accomplished. Wright's F-statistics values obtained from the results of global AMOVA revealed 11.7% deficit of heterozygotes for each of the analyzed breeds (F_{IS}) whereas the total population had a 27.8% deficit of heterozygotes (F_{IT}). The average genetic differentiation (F_{ST}) between the breeds was 18.2% ($P<0.05$) indicating significantly higher discrimination between breeds (Table 4). Details of AMOVA results are presented in Table 4. The pair-wise F_{ST}, Slatkin linearized F_{ST}, and Nei's distance (d) values were used to illustrate the genetic distance between breeds (Fig. 2a, 2b & 2c), which significantly differentiated all three breeds.

Murrah and Nili-Ravi population were clustered together while the Gojri population was present as a distinct group, suggesting it as a different breed in factorial correspondence analysis (Fig. 3) and phylogenetic tree (Fig. S2).

Effective population size (N_e) was estimated excluding rare alleles with an allele frequency below 0.05. The estimated effective population size of Gojri, Nili-Ravi, and, Murrah was found to be 142, 75, and 556, respectively. The Jack-knife CIs for the Ne estimates were 83-396, 48-141, and 136 to infinity for Gojri, Nili-Ravi, and Murrah, respectively at 0.05 *P*-critical value of rare alleles.

Bayesian genetic structure

Number of possible sub-populations estimated through Evanno Δ*K* method suggested a maximum of three populations (Fig. 4). Population assignment accomplished in STRUCTURE for K=2, 3, and 4 and results are presented in the form of bar plot (Fig. 5). For K=3, as estimated through Evanno Δ*K* method, it showed 99.4% of Gojri buffaloes are classified into their pre-defined breed. 95.9% of Nili-Ravi and 83.6% of Murrah were assigned to their respective pre-defined groups.

Inferred ancestry of each individual (for *K*=3) along with average proportion of each individuals classified into respective pre-assigned breeds (for *K*=2, 3, and 4) is reported (Table 5).

Discussion

In India, limited work on complete characterization and classification of buffalo genetic resources have been carried out in past, primarily due to availability of much acclaimed Murrah buffaloes. The native breeding tract of Murrah buffalo is North India, and currently, more than 40% of the countries buffalo population is either Murrah or has been crossed with Murrah buffaloes. Hence, genetic studies on other buffalo populations is often neglected. However, several studies have been taken up on morphometric characterization of individual breeds yet there are limited reports on genetic diversity studies through molecular markers and comparative studies. A parallel approach of characterization and classification of buffalo germplasm in a region is much needed for genetic improvement in such populations. The present study is an evaluation of Riverine buffaloes of North India taking a geographical region-based approach.

Morphological diversity

Gojri animals with unique phenotypic appearance are quite distinct from Murrah, Murrah crosses, and Nili Ravi (Vohra et al., 2012). The average measurements for body biometric traits across the studied buffalo populations of the North India is listed in Table 1. Thirteen body biometric traits across 3 population when compared, revealed significant differences among the studied populations, except for FL and EL among Murrah and Gojri buffaloes, HC and HL between Nili-Ravi and Gojri population. Body height (HT) and face width (FW) did not vary among Murrah and Nili-Ravi populations. The comparison of morphometric traits between all three buffalo breeds of the North India outlined the phenotypic distinctness for majority of the body biometric trait. The

coefficient of variation (CV) percentage was least for body height in all three breeds. On comparing average of HT, BL, CG, and PG, Gojri buffaloes were found to be of smaller size than Murrah and Nili-Ravi. Nivsarkar et al. (2000) in Nili-Ravi reported average HT, CG, and BL as 134.2, 207.7, and 165.4 cm, respectively, which is comparable to our results. CV% was highest for HL in Gojri (19.52%) and Murrah (12.61%) buffaloes indicating lesser selection pressure on them and more environmental influence. Face width (FW) was least variable in Murrah and Nili-Ravi while it varied greatly in Gojri buffaloes. Most of the body biometric traits measured were less variable indicating their reliability in population classification studies.

In the canonical discriminant analysis, two functions were needed for separation of three distinct population (Asamoah-Boaheng and Sam, 2016) and the first function (function 1) explains 66.7% of the variance and has a Wilk's lambda (0.008) with $p < 0.05$. The second function explains only 33.3% of the variance in the data, with a recorded $p < 0.05$ for Wilk's lambda (0.122). Wilks' Lambda value close to zero represents a greater number of variables contribute to the discriminant function (Toalombo Vargas et al., 2019), thus the first function in this study plays major role in classifying the breeds.

Microsatellite variations and Genetic diversity

Microsatellite marker data being the best-suited molecular information for the assessment of genetic diversity (Bowcock et al., 1994; Laval et al., 2000; Groeneveld et al., 2010), allows future management and conservation of the breeds based on their genetic architecture (Luikart et al., 2003; Taberlet, et al., 2008; Toro et al., 2009; Teneva et al., 2013). The FAO and the ISAG/FAO Advisory Group on Animal Genetic Diversity have proposed a panel of 25 SSR markers for

diversity studies in buffaloes (Singh et al., 2018). Hence, in the present study the 22 highly polymorphic microsatellite markers out of 25 marker panel, were used for diversity analysis.

The mean number of alleles (N_a) in population over a range of loci is considered a fair indicator of allelic variation. The mean N_a ranged from 0-10, 0-13, and 3-14 in Nili-Ravi, Murrah, and Gojri buffaloes, respectively. The mean N_a per locus for each population in the present study is similar to the reports of Kathiravan et al. (2010) in South Kanara buffaloes; Marques et al. (2011) in Brazilian buffaloes, Martinez et al. (2006), Bhuyan et al. (2010) in Murrah buffaloes and Sajid Ali et al. (2020) in Purnathadi buffaloes. However, a higher number of alleles per locus ranged from 11-26 alleles in Indian water buffaloes have been reported by Vijh et al. (2008). The type of breed under investigation, usage of the particular panel of microsatellite markers, methods of genotyping and the genetic polymorphism within the breed itself greatly influence this variation in the N_a.

For microsatellite data, Ohta and Kimura (1973) have established the relationship between the expected homozygosity and its estimator θ, under a pure stepwise mutation model *i.e.* expected homozygosity = $1/\sqrt{1+2\theta}$. An estimator of θ can be obtained from microsatellite data by applying the formula, $\theta_H = [1/ (1-H_e)^2-1]$ (Excoffier et al., 2010), where H_e is the expected heterozygosity. The mean H_e ranging from 0.14 to 0.81 across all three population over all loci is indicative of sufficient polymorphism to measure genetic variation (Takezaki and Nei, 1996). In Gojri buffaloes the expected heterozygosity (H_e) ranged from 0.12 to 0.82 that was comparable with the results reported by Singh et al. (2019). While in both Murrah and Nili-Ravi, it ranged from 0 to 0.81. Similar high overall mean H_e were reported in Pandharpuri (Khade et al., 2019), Mehsana (Jakhesara et al., 2010), Egyptian (Attia et al., 2014) and Purnathadi (Sajid Ali et al., 2020) buffaloes. The substantially high H_e values implies the presence of high genetic variability in the studied buffalo breeds and suitability of the marker panel for the present study.

The average F statistics over 19 loci were F_{IS} = 0.11744, F_{ST} = 0.18252 and F_{IT} = 0.27852, In the present study considerable degree of differentiation has been estimated compared to other buffalo populations from different regions, probably because these populations are genetically distinct. Joshi et al. (2012) reported an F_{ST} value of 7.2% in buffaloes of Indo-gangetic plain, while Vijh et al. (2008) reported a value of 9.69%. However, a comparatively lesser value in eight Indian riverine buffalo was reported by Kumar et al. (2006) which was 3.4%. This value suggested the existence of greater genetic differentiation among North-Indian buffalo breeds than breeds found all over India. A heterozygote deficiency was evident from the positive mean F_{IS} value (0.117>0) indicating low to moderate amount of inbreeding in the population. This could be attributed to assortative mating in small herds owned by farmers, genetic hitchhiking, or the null alleles (Mishra et al., 2008). However, AMOVA over all 22 loci showed 23.59% of variations between populations suggesting the distinctness of all three breeds. The F_{IS} value was found to be 4.74%, which is comparable to values obtained in Purnathadi buffaloes (Sajid Ali et al., 2020).

The pair-wise F_{ST} values ranged from 0.09 between Murrah and Nili-Ravi to 0.32 between Nili-Ravi and Gojri breeds. The F_{ST} between Murrah and Gojri was 0.25 (Fig. 2a). Least differentiation was found between Nili-Ravi and Murrah (0.09) based on Slatkin linearized F_{ST} while it was highest between Nili-Ravi and Gojri (0.46). Between Murrah and Gojri it was found to be 0.33 (Fig. 2b). Nei's distance (d); average within and between populations differentiation is presented in the form of a heat map (Fig. 2c), that shows least distance between Murrah and Nili-Ravi breeds and discriminate Gojri as another population. These results were also in compliance with the results from the factorial correspondence analysis based on molecular data and phylogenetic tree obtained from dissimilarity matrix. In the scatter plot of factorial analysis, Murrah and Nili-Ravi are

278 invariably clustered together. Meanwhile, Gojri was found to be plotted on the opposite side of
279 axis-2 (Fig. 3), yet with more scattering among individuals.

280 The linkage disequilibrium method relies on measures of departure from expected genotype and
281 gametic frequencies, which is the basis for estimation of effective population size (Hill, 1981;
282 Waples, 1991; Luikart et al., 2010). The N_e estimated from microsatellite data reflects the true
283 population distribution of North Indian buffaloes. The comparatively higher N_e of Murrah
284 buffaloes is due to the larger population distribution of the breed in India. The N_e estimates of Gojri
285 population reflects its present status and probable serious inbreeding in future. Hence, the ongoing
286 indiscriminate breeding practices should be shifted to implementation of organized breeding
287 policies focussed on conservation of this distinct breed.

288 **Bayesian genetic structure**

289 Structure software (Pritchard et al., 2000) was used to determine the unbiased structure assuming
290 no prior knowledge regarding the number of breeds. The highest delta K (ΔK) value was calculated
291 as previously described (Evanno et al., 2005). The optimum ΔK value (Fig. 4), was found at K =
292 3. For K=2, there was no differentiation between Nili-Ravi and Murrah breed. One individual from
293 Murrah population showed significant level of admixture from Gojri population. 99.7% of Gojri
294 buffaloes were classified as a different breed whereas 99.7% and 98.9% of Nili-Ravi and Murrah
295 buffaloes were assigned to one single population, respectively. When K is assumed to be four,
296 Gojri buffaloes are assigned to one distinct cluster with 99% of memberships. Structure assigned
297 all three population into three different breeds when K was assumed to be 3. This indicated that
298 studied populations has well differentiated and possess unique allelic combinations despite being
299 reared in similar geographical regions. However, a low to moderate amount of admixture could be

observed in both Murrah and Nili-Ravi population. For *K*=3, Nili-Ravi showed an average admixture of 3.7% from Murrah and 0.4% from Gojri buffaloes while it was quite high for Murrah with an average admixture of 15.2 and 1.2% from Nili-Ravi and Gojri, respectively. While Gojri population was found to have 99.4% pure blood with an admixture of 0.3% from Nili-Ravi and 0.4% from Murrah. Our results indicate the presence of sufficiently large genetic variability among the North Indian Riverine buffaloes. However, Gojri buffalo populations is unique, compared to Murrah and Nili Ravi buffalos, which were found to be genetically closer than expected. Presently the breeding areas of all these populations are overlapping due to adoption of Murrah as an improver breed for milk production, thus leading to its dominance over Nili-Ravi and Gojri buffalo.

Conclusion

This study demonstrated that the characterization and classification of genetic diversity in Indian buffaloes could be better accomplished through a parallel approach comprising morphometric traits and microsatellite markers. Study of buffalo genetic diversity of Northern India revealed admixture of two major dairy buffalo breeds and a distinct buffalo population was identified. The results obtained provides an opportunity for the design of genetic improvement programs with appropriate choice of breeds for upgrading local non-descript buffaloes along with conservation of unique germplasm. The estimates of effective population size and fixation indices indicate absence of intense systematic selection in past. Further studies involving large populations including samples from other regions of Indian buffalo with FAO recommended microsatellite loci are required to understand the genetic relationships among buffalo genetic resource of India.

Acknowledgement

The authors thank the Directors of ICAR-NDRI and ICAR-NBAGR for their support to carry out the study. Technical support received from Mr. Subhash Chander, T-5, NBAGR, is gratefully appreciated.

References:

1. Ali, S. S., Kuralkar, S. V., Das, R., Raina, V., Kataria, R. S. & Vohra, V. (2020). Assessment of genetic diversity and bottleneck in Purnathadi buffaloes using short tandem repeat markers. *Animal Biotechnology*, 1-12.
2. Asamoah-Boaheng, M. & Sam, E. K. (2016). Morphological characterization of breeds of sheep: a discriminant analysis approach. *SpringerPlus* **5**(1), 69.
3. Attia, M., Abou-Bakr, S. & Hafez, Y. M. Genetic polymorphism of seven microsatellite DNA markers in Egyptian buffalo. *Animal Biotechnology* (Cattles, Buffalo) **7** (12), 7.
4. Barker, J. S. F. (1999). Conservation of livestock breed diversity. *Animal Genetic Resources/Resources génétiques animales/Recursos genéticos animals* **25**, 33-43.
5. Bhuyan, D.K., Sangwan, M.L., Gole, V.C. & Sethi, R.K. (2010). Studies on DNA fingerprinting in Murrah buffaloes using microsatellite markers. http://nopr.niscair.res.in/handle/123456789/10433
6. Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R. & Cavalli-Sforza, L. L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**(6470), 455-457.
7. Breno Araújo de Melo, Isabele de Melo Nascimento, Lays Thayse Alves dos Santos, Luciano Gomes de Lima, Filipe Chagas Teodózio de Araújo, Raisa Rodrigues Santos Rios, Alberto de Gusmão Couto & Angelina Bossi Fraga (2018). Body morphometric

measurements in Murrah crossbred buffaloes (Bubalus bubalis). *Journal of Applied Animal Research* **46**:1. 1307-1312. DOI: 10.1080/09712119.2018.1502669

8. Bruford, M. W., Bradley, D. G. & Luikart, G. (2003). DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics* **4**(11), 900.
9. Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J. & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular ecology resources* **14**(1), 209-214.
10. Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology* **14**(8), 2611-2620.
11. Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources* **10**(3), 564-567.
12. FAO (2011). Molecular genetic characterization of animal genetic resources. In: FAO Animal Production and Health Guidelines. No. 9, Rome.
13. FAO (2012). Phenotypic characterization of animal genetic resources. In: FAO Animal Production and Health Guidelines No. 11, Rome. Available at: http://www.fao.org/3/a-i2686e.pdf
14. Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Molecular ecology resources*, **17**(1), 27-32.

15. Groeneveld, L.F., Lenstra, J.A., Eding, H., Toro, M.A., Scherf, B., Pilling, D., Negrini, R., Jianlin, H., Finlay, E.K., Groeneveld, E. & Weigend S. (2010). Genetic diversity in livestock breeds. *Animal Genetics* **41**(Suppl 1), 6-31.

16. Hall, S. J. & Bradley, D. G. (1995). Conserving livestock breed biodiversity. *Trends in ecology & evolution* **10**(7), 267-270.

17. Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetics Research* **38**(3), 209-216.

18. Jakhesara, S. J., Rank, D. N., Kansara, J. D., Parikh, R. C., Vataliya, P. H. & Solanki, J. V. (2010). Microsatellite DNA typing for assessment of genetic variability in the Mehsana buffalo breed of India. *BUFFALO BULLEITN IBIC, KASETSART UNIVERSITY, PO BOX 1084 BANGKOK 10903, THAILAND* **29**(4), 262.

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

20. Johnson, R. A. & Wichern, D. W. (2002). *Applied multivariate statistical analysis*, Upper Saddle River, NJ: Prentice hall **5**(8).

21. Kathiravan, P., Mishra, B. P., Kataria, R. S., Goyal, S., Tripathy, K. & Sadana, D. K. (2010). Short tandem repeat based analysis of genetic variability in Kanarese buffalo of South India. *Russian journal of genetics* **46**(8), 988-993.

22. Khade, K. A., Panigrahi, M., Ahmad, S. F., Kumar, P. & Bhushan, B. (2019). Genetic characterization and assessment of diversity in Pandharpuri buffalo breed of India using heterologous microsatellite markers. *Animal biotechnology*, 1-6.

387 23. Kumar, S., Gupta, J., Kumar, N., Dikshit, K., Navani, N., Jain, P. & Nagarajan, M. (2006).
388 Genetic variation and relationships among eight Indian riverine buffalo breeds. *Molecular*
389 *ecology* **15**(3), 593-600.
390 24. Laval, G., Iannuccelli, N., Legault, C., Milan, D., Groenen, M. A., Giuffra, E. &
391 Geldermann, H. (2000). Genetic diversity of eleven European pig breeds. *Genetics*
392 *selection evolution* **32**(2), 187.
393 25. Li, M. H., Nogovitsina, E., Ivanova, Z., Erhardt, G., Vilkki, J., Popov, R. & Kantanen, J.
394 (2005). Genetic contribution of indigenous Yakutian cattle to two hybrid populations,
395 revealed by microsatellite variation. *Asian-australasian journal of animal sciences* **18**(5),
396 613-619.
397 26. Luikart, G., England, P. R., Tallmon, D., Jordan, S. & Taberlet, P. (2003). The power and
398 promise of population genomics: from genotyping to genome typing. *Nature reviews*
399 *genetics* **4**(12), 981-994.
400 27. Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K. & Allendorf, F. W. (2010).
401 Estimation of census and effective population sizes: the increasing usefulness of DNA-
402 based approaches. *Conservation Genetics* **11**(2), 355-373.
403 28. Marques, J. R. F., Martínez, A. M., Costa, M. R., Albuquerque, M. S. M., Quiroz, J., Plá,
404 J. L. V. & Bermejo, J. V. D. (2011). Genetic diversity of Brazilian buffaloes (Bubalus
405 bubalis) using DNA microsatellites. *Archivos de zootecnia* **60**(232), 1213-1221.
406 29. Martínez, E., Tirado, J. F., Cerón-Muñoz, M. F., Moreno, M., Montoya, A., Corrales, J. D.
407 & Calvo, S. J. (2009). Genetic characterization of Murrah Buffalo breed in Colombia using
408 microsatellite DNA markers. *Livestock Research for Rural Development* **21**(1).

30. Metta, M., Kanginakudru, S., Gudiseva, N. & Nagaraju, J. (2004). Genetic characterization of the Indian cattle breeds, Ongole and Deoni (Bos indicus), using microsatellite markers–a preliminary study. *BMC genetics* **5**(1), 16.

31. Ministry of Agriculture and Farmers Welfare, GoI. 19th Livestock Census Report. New Delhi, India: Department of Animal Husbandry DAHDF; 2012.

32. Mishra, B. P., Kataria, R. S., Bulandi, S. S., Kumar, V. & Mukesh, M. (2008). Genetic diversity in river buffalo (Bubalus bubalis) breeds of central India using heterologous bovine microsatellite markers. *Journal of Applied Animal Research* **33**(2), 159-163.

33. Mishra, B.P., Kataria, R.S., Kathiravan, P., Singh, K.P., Sadana, D.K. & Joshi, B.K. (2010). Microsatellite based genetic structuring reveals unique identity of Banni among river buffaloes of Western India. *Livestock Science* **127**(2–3), 257-261

34. Nei, M. (1987). *Molecular evolutionary genetics*. Columbia university press, New York, NY, USA.

35. Nivsarkar, A.E., Vij, P.K. & Tantia, M.S. (2000). Animal genetic resources of India: Cattle and Buffalo. Directorate of information and publications of Agriculture, ICAR, New Delhi, India.

36. Ohta, T. & Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetics Research* **22**(2), 201-204.

37. Pandey, A. K., Sharma, R., Singh, Y., Prakash, B. B. & Ahlawat, S. P. S. (2006). Genetic diversity studies of Kherigarh cattle based on microsatellite markers. *Journal of genetics* **85**(2), 117-122.

38. Pandey, A. K., Sharma, R., Singh, Y., Prakash, B. & Ahlawat, S. P. S. (2006). Evaluation of genetic variability in Kenkatha cattle by microsatellite markers. *Asian-australasian journal of animal sciences* **19**(12), 1685-1690.

39. Perrier, X., Flori, A. & Bonnot, F. (2003). Data analysis methods In: Hamon P, Seguin M, Perrier X, Glaszmann JC Ed, Genetic diversity of cultivated tropical plants.

40. Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**(2), 945-959.

41. Ruane, J. (2000). A framework for prioritizing domestic animal breeds for conservation purposes at the national level: a Norwegian case study. *Conservation Biology* **14**(5), 1385-1393.

42. Sambrook, J. & Russel, D.W. (2001). Molecular Cloning, Animal Laboratory manual. Cold Spring Harbor Laboratory press, Cold spring Harbor, New York.

43. Sharma, R., Maitra, A., Singh, P. K. & Tantia, M. S. (2013). Genetic diversity and relationship of cattle populations of East India: distinguishing lesser known cattle populations and established breeds based on STR markers. *SpringerPlus* **2**(1), 359.

44. Simianer, H. (2005). Decision making in livestock conservation. *Ecological Economics* **53**(4), 559-572.

45. 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.

46. 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.

454 47. Sodhi, M., Mukesh, M., Mishra, B. P., Prakash, B., Ahlawat, S. P. S. & Mitkari, K. R.
455 (2005). Evaluation of genetic differentiation in Bos indicus cattle breeds from Marathwada
456 region of India using microsatellite polymorphism. *Animal biotechnology* **16**(2), 127-137.
457 48. SPSS. 2001. Statistical Package for Social Sciences. SPSS Inc., 444 Michigan Avenue,
458 Chicago, IL 60611.
459 49. Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite
460 populations. *Genetics* **105**(2), 437-460.
461 50. Tajima, F. (1993). Measurement of DNA polymorphism. *Mechanisms of Molecular*
462 *Evolution, In: Introduction to Molecular Paleopopulation Biology*, 37-59.
463 51. Takezaki, N. & Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees
464 from microsatellite DNA. *Genetics*, **144**(1), 389-399.
465 52. Tantia, M. & Vijh, R. (2012). Microsatellite analysis of buffaloes of Indo-Gangetic
466 Plains. *Indian Journal of Animal Sciences* **82**(11), 1434-1437.
467 53. Teneva, A., Dimitrov, K., Caro Petrović, V., Petrović, M. P., Dimitrova, I., Tyufekchiev,
468 N. & Petrov, N. (2013). Molecular genetics and SSR markers as a new practice in farm
469 animal genomic analysis for breeding and control of disease disorders. *Biotechnology in*
470 *Animal Husbandry* **29**(3), 405-429.
471 54. Toalombo Vargas, P. A., León, J. M., Fiallos Ortega, L. R., Martinez, A., Villafuerte
472 Gavilanes, A. A., Delgado, J. V. & Landi, V. (2019). Deciphering the Patterns of Genetic
473 Admixture and Diversity in the Ecuadorian Creole Chicken. *Animals* **9**(9), 670.
474 55. Toro, M. A., Fernández, J. & Caballero, A. (2009). Molecular characterization of breeds
475 and its use in conservation. *Livestock Science* **120**(3), 174-195.

476 56. Toro, M. & Caballero, A. (2005). Characterization and conservation of genetic diversity in
477 subdivided populations. Philosophical Transactions of the Royal Society of London, Series
478 B: Biological Sciences **360**, 1367–78.
479 57. Vijh, R. K., Tantia, M. S., Mishra, B. & Bharani Kumar, S. T. (2008). Genetic relationship
480 and diversity analysis of Indian water buffalo (Bubalus bubalis). *Journal of animal science*
481 **86**(7), 1495-1502.
482 58. Vohra, V., Sodhi, M., Niranjan, S. K., Mishra, A. K., Chopra, A., Kumar, M. and Joshi A
483 K. (2017) Characterization of rare migratory cattle and evaluation of its phylogeny using
484 short-tandem-repeat-based markers. *Journal of Applied Animal Research* **45**(1): 355-363
485 59. Vohra, V., Niranjan, S. K. & Joshi, B. K. (2012). Gojri–A novel migratory buffalo
486 germplasm in Punjab and Himachal Pradesh. *Journal of Animal Research* **2**(3), 317-321.
487 60. Vohra, V., Niranjan, S. K., Mishra, A. K., Jamuna, V., Chopra, A., Sharma, N. & Jeong,
488 D. K. (2015). Phenotypic characterization and multivariate analysis to explain body
489 conformation in lesser known buffalo (Bubalus bubalis) from North India. *Asian-*
490 *Australasian journal of animal sciences* **28**(3), 311.
491 61. Waples, R. S. (1991). Genetic methods for estimating the effective size of cetacean
492 populations. *Report of the International Whaling Commission (special issue)* **13**, 279-300.
493 62. Weitzman, M. L. (1993). What to preserve? An application of diversity theory to crane
494 conservation. *The Quarterly Journal of Economics* **108**(1), 157-183.
495 63. Yadav, D. K., Arora, R. & Jain, A. (2017). Classification and conservation priority of five
496 Deccani sheep ecotypes of Maharashtra, India. *PloS one* **12**(9).
497 64. Yoon, D. H., Kong, H. S., Oh, J. D., Lee, J. H., Cho, B. W., Kim, J. D. & Lee, H. K. (2005).
498 Establishment of an individual identification system based on microsatellite

499 polymorphisms in Korean cattle (Hanwoo). *Asian-australasian journal of animal*
500 *sciences* **18**(6), 762-766.
501 65. Zouros, E. (1979). Mutation rates, population sizes and amounts of electrophoretic
502 variation of enzyme loci in natural populations. *Genetics* **92**(2), 623-646.

503

504

Table 1(on next page)

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.

1

2

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

Table 2(on next page)

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	0.000
2nd function	7.19	33.3	100	0.937	0.122	493.30	9	0.000

1

Table 3(on next page)

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

Locus	Expected Heterozygosity (H_e)			Theta H (θ_H)			Number of alleles (N_a)		
	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
CSSM19	0.74	0	0.73	2.80	0	2.75	6	0	6
CSSM33	0.71	0.73	0.63	2.49	2.67	1.74	8	9	7
CSSM45	0.65	0.80	0.73	1.87	3.98	2.72	5	6	6
CSSM47	0.81	0.69	0.82	4.23	2.25	4.63	10	10	11
CSSM66	0.79	0.61	0.82	3.67	1.60	4.43	7	6	9
Hel013	0.67	0.75	0.82	2.08	3	4.49	8	8	9
ILSTS19	0.14	0.17	0.12	0.16	0.20	0.14	2	6	3
ILSTS25	0.58	0.61	0.72	1.39	1.59	2.52	5	6	8
ILSTS26	0.67	0.61	0.76	2.05	1.57	3.26	6	6	5
ILSTS28	0.76	0.76	0.76	3.25	3.16	3.21	8	7	6
ILSTS29	0.33	0.26	0.82	0.49	0.35	4.49	6	4	10
ILSTS30	0.71	0.60	0.69	2.45	1.47	2.25	7	6	6
ILSTS33	0.66	0.68	0.59	1.99	2.14	1.47	6	7	3
ILSTS36	0.66	0.67	0.72	1.98	2.03	2.58	6	6	12
ILSTS52	0.67	0	0.73	2.00	0	2.72	9	0	8
ILSTS56	0.40	0.53	0.59	0.67	1.12	1.45	5	7	8
ILSTS58	0.81	0.85	0.77	4.24	5.70	3.40	7	9	11
ILSTS60	0.45	0.36	0.60	0.81	0.57	1.53	4	3	14
ILSTS61	0.66	0.81	0.76	1.91	4.23	3.23	6	13	11
ILSTS089	0	0.79	0.76	0	3.76	3.14	0	6	6
ILSTS95	0.80	0.71	0.71	4.08	2.43	2.48	10	11	8
Mean	0.61	0.58	0.70	1.55	1.36	2.33	6.27	6.59	7.86
SD	0.22	0.25	0.15	-	-	-	2.33	3.10	2.83

1

Table 4(on next page)

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})	0.0000
Among individuals within populations	125	786.65	0.73	9.60	0.117 (F_{IS})	0.0000
Within individuals	128	640.50	5.46	72.15	0.278 (F_{IT})	0.0000
Total	255	1659.37	7.57	100	-	-

1

Table 5(on next page)

Individual wise ancestry level inferred through Bayesian method in Structure

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

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

1

Figure 1

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

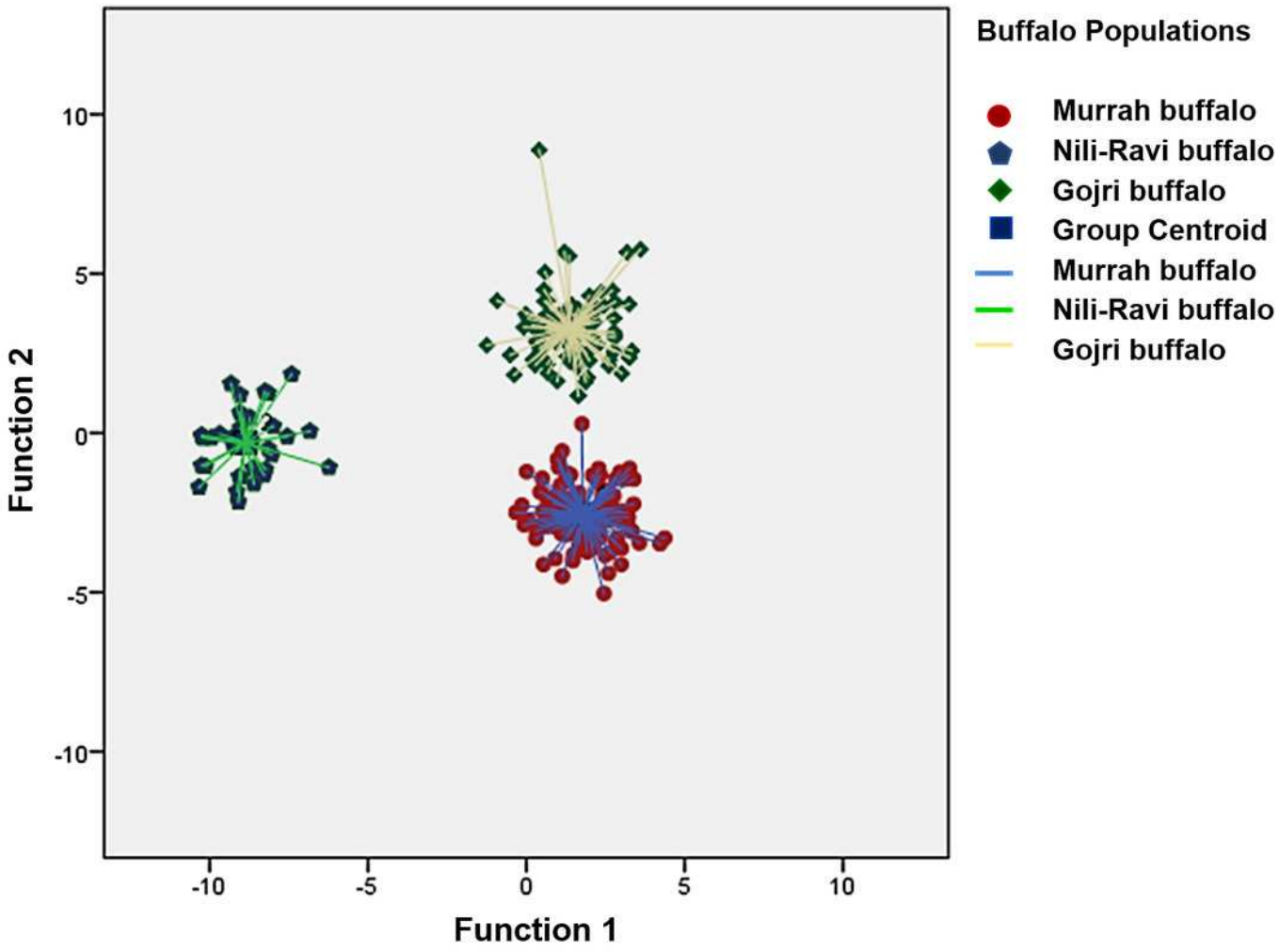

Figure 2

Heatmap of molecular diversity indices

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

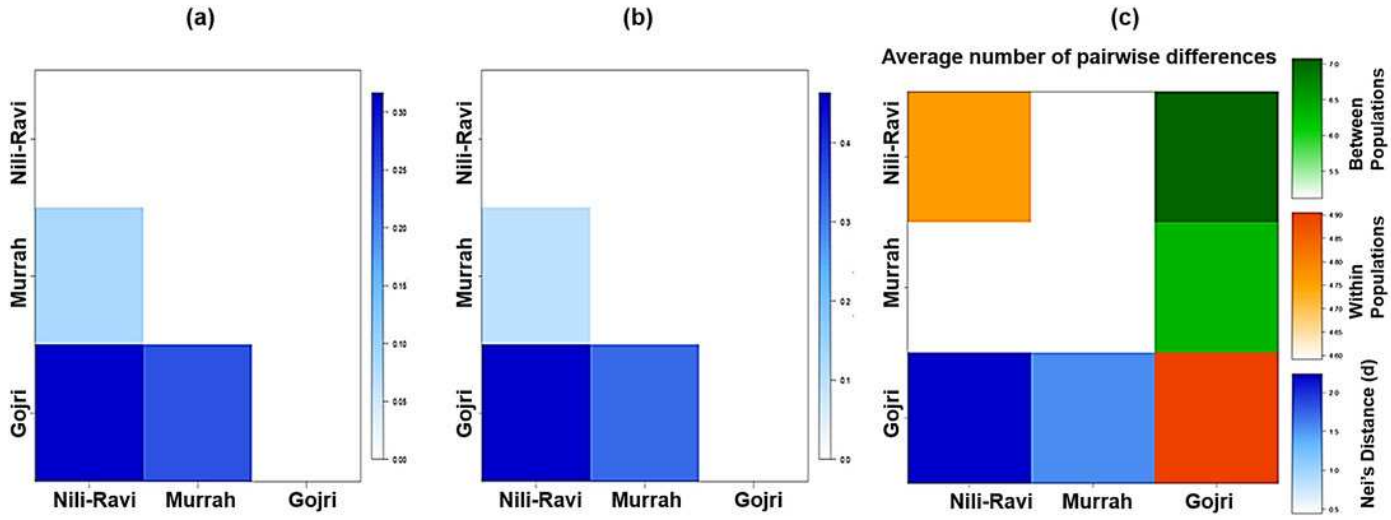

Figure 3

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

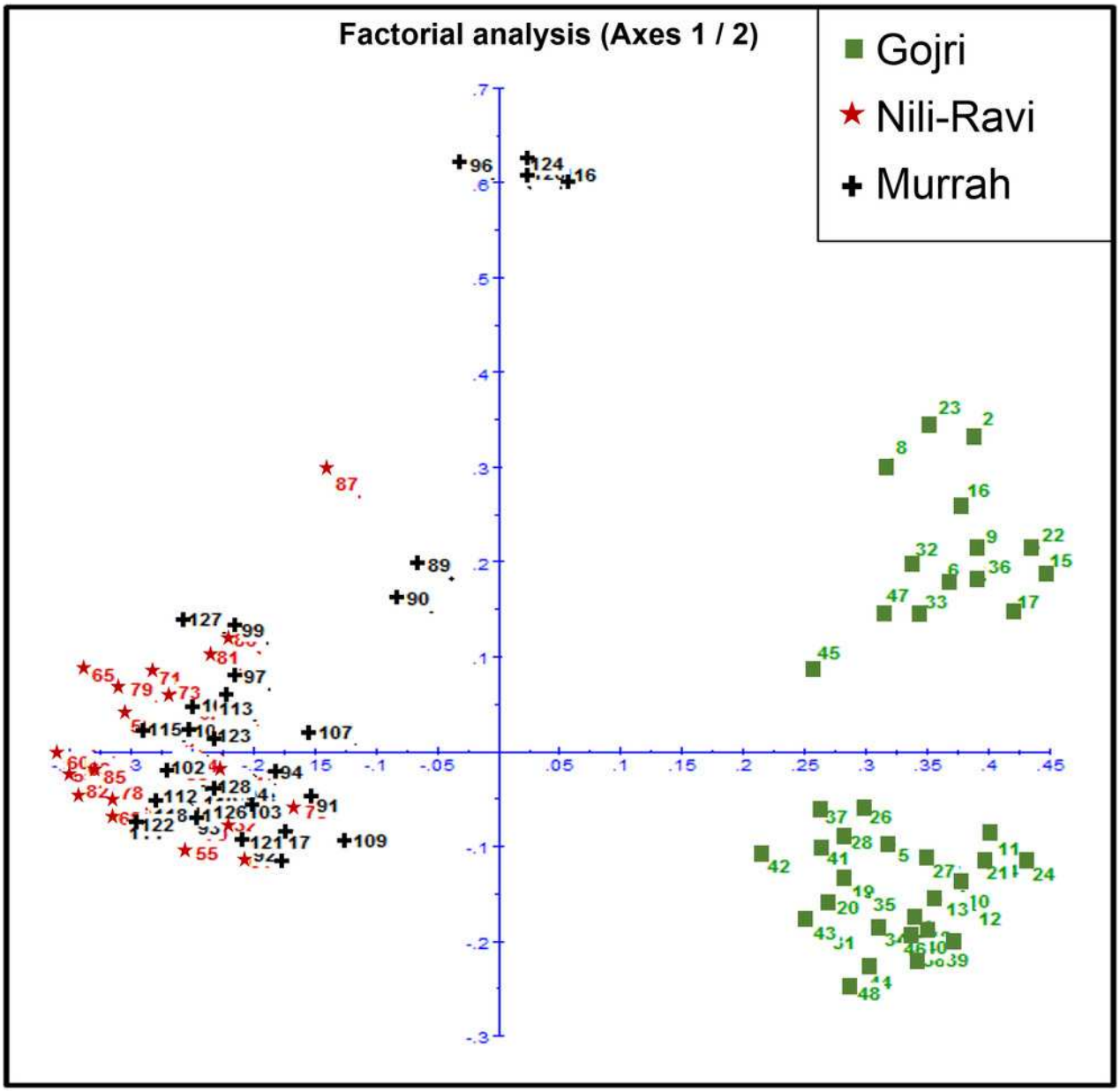

Figure 4

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

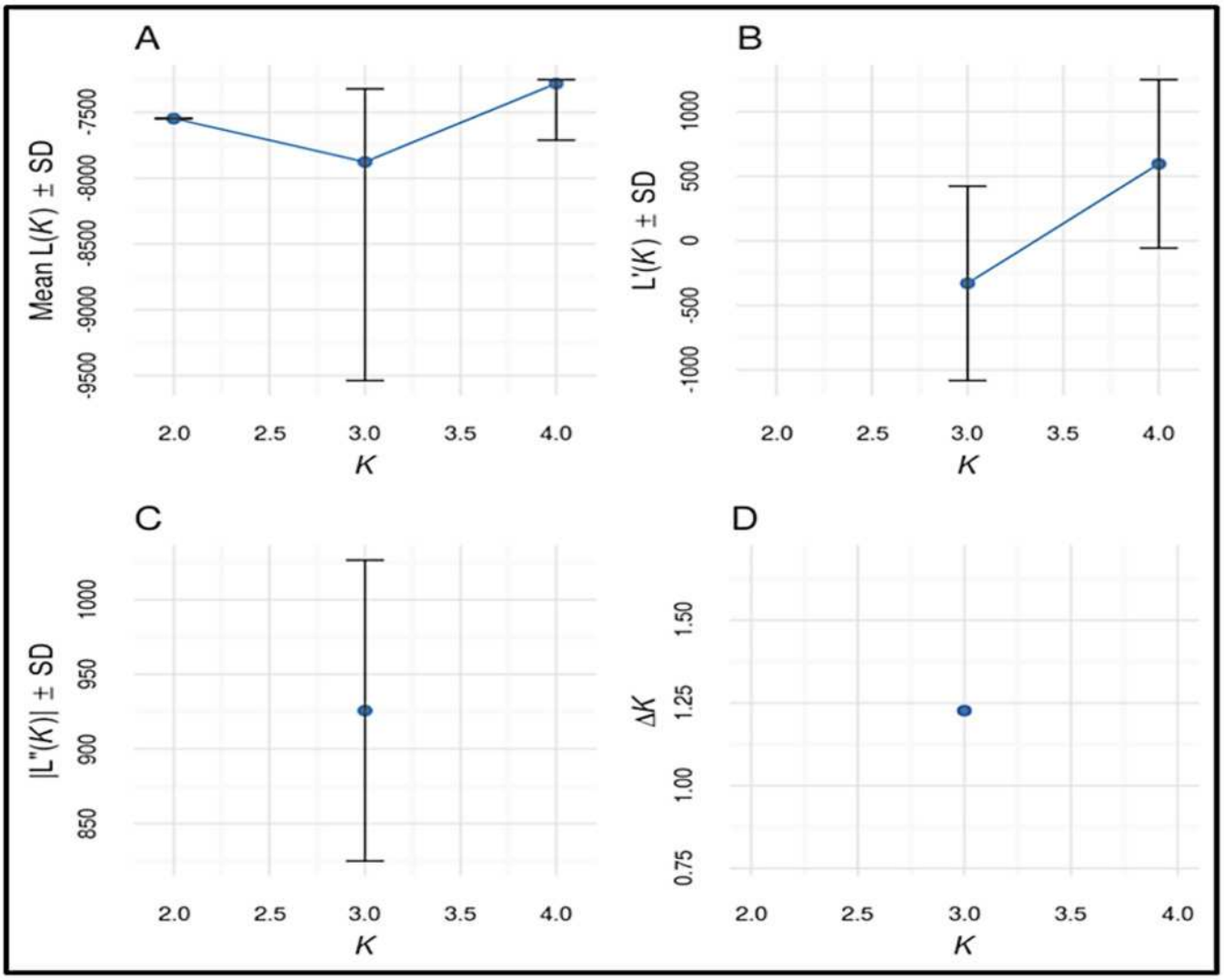

Figure 5

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).