Indian Journal of Animal Sciences **81** (8): 836–848, August 2011

Genetic landscape and demography of buffaloes in Indo- Gangetic plains

S UPASNA[1], JYOTI JOSHI[2], PRIYANKA BANERJEE[3] and R K VIJH[4]

National Bureau of Animal Genetic Resources, Karnal, Haryana 1320 01 India

Received: 9 March 2011; Accepted: 29 March 2011

ABSTRACT

The Indo-Gangetic plains have about a quarter of the total buffalo population in the country, yet there have been only one defined breed of buffalo in this vast plains traversed by 2 major rivers of the country and their large number of tributaries. We generated data on 625 buffaloes using 11 microsatellite loci and carried out the statistical analysis to reveal genetic landscape, demographic parameters of these buffaloes and to investigate the existence of genetic structures underlying the continuity of geographical landscape. The investigations revealed that there is isolation by distance and existence of 5 genetic structures, though these structures do not have continuity among the sampled areas. The analysis of data on buffaloes of Indo-Gangetic plains revealed that there has not been any recent colonization event nor severe reduction in the effective population size. There has been a historical constancy of size of buffalo in this geographical area as revealed by k and g tests. The analysis revealed aggregation of alleles pointing towards absence of randomness in the geographical landscape. The Moran and Geary's index also reveal non randomness of the distribution of allele pointing towards existence of population structure in the Indo-Gangetic Plains. The analysis of variance revealed 6% variation attributable to districts component. The existence of major rivers and their tributaries do not have significant effect on the structuring of the populations as revealed by partial Mantel tests.

Key words: Allelic aggregation index, Buffaloes, Demography, Genetic landscape, Isolation by distance, Spatial auto-correlation

The buffalo population of Uttar Pradesh constitutes 23.4% of the total buffalo population of India, yet this population has not been classified into distinct breeds or populations. The Indo- Gangetic plains have only one defined breed of buffalo named Bhadawari. There is large number of rivers traversing the Indo-Gangetic plains. The major ones are Ganga and Yamuna and their tributaries. The rivers can be one of the potential causes for the structuring of the buffalo populations. The populations are said to show genetic structure whenever the distributions of their genes do not confirm to panmictic expectations. The interaction between landscape features and micro-evolutionary processes, such as gene flow and genetic drift may play an important role in the structuring of buffalo populations. In this study, we have made an attempt to analyse the microsatellite data and geographic landscape to arrive at population genetic parameters which help in inferring the demographic features of a given population ultimately leading to infer genetic structures if they exist and help in taking informed conservation decisions.

Present address: [1]Research Associate, [2,3]Senior Research Fellow, [4]Principal Scientist (e mail: rameshkvijh@gmail.com).

MATERIALS AND METHODS

Blood samples of 625 buffaloes were collected from throughout the Indo-Gangetic plains. The districts with very few samples were clubbed with the adjoining districts and a total of 34 conglomerates. The DNA was isolated following normal protocols (Sambrook *et al*. 1989) and 11 heterologous microsatellites were selected from cattle database. The microsatellite data were generated using fluorescently labeled primers and using automated DNA sequencer. The data generated were extracted using GeneMapper software version 4.1. The Slatkin's F_{ST} (Slatkin and Voelm 1991) was estimated using the software Arlequin.

The allelic aggregation index was calculated using the software Alleles In Space (AIS) (Miller 2005). The allelic aggregation index analysis on genotypes rather than individual alleles was performed using AIS. This avoids a situation in which the nearest neighbor distance for an instance of an allele in a homozygous individuals will be 0, indicating that the procedure may be sensitive to departures from Hardy-Weinberg genotypic proportions that could exist in a natural population. The area was estimated from the default option of the software which is defined as the area encompassed by the rectangle defined by the maximum and

minimum coordinates provided in the data sets coordinate file. The area was estimated from the polygonal estimate provided in the "Plots of Sample Location" feature contained in AIS. In addition to performing allele specific tests, we also performed a global test by calculating *RjAVE* over alleles and loci. As with the allele specific tests, P-values for the global analyses were obtained through random allocation of individuals and genotypes over the specific set of coordinates sampled.

For evaluation of the microsatellite data for the phylogeographic analysis, we utilized landscape shape interpolation of AIS software. This procedure revealed visualizations of patterns of diversity across landscape. It produces a 3-dimensional surface plot where X and Y axes correspond to geographical locations and surface heights (Z-axis) represent genetic distances. We constructed a connectivity network among all of the sampling locations in the data set. We utilised "all pair-wise locations" connectivity network (Fig 1) based on Delaunay triangulations (Watson 1992, Brouns *et al*. 2003).

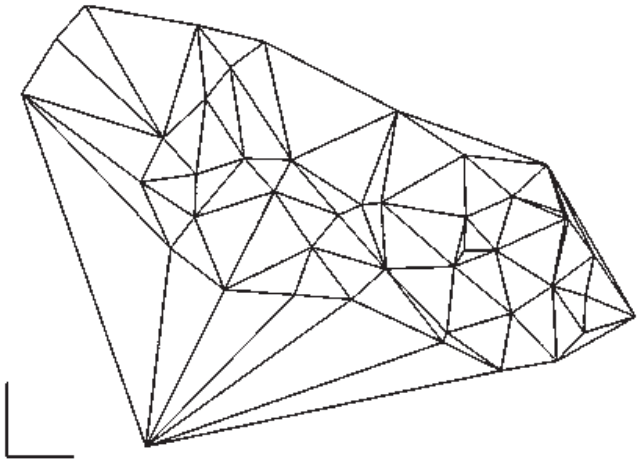

Fig. 1. Connectivity network based on Delaunay triangulations

In the present study, we used residual genetic distances or "pseudo slopes" for the analysis. The genetic distance was calculated as:

$$D_{ij} = \frac{\sum_{k=1}^{n}\left(1-\sum_{l=1}^{q_k}\sqrt{P_{il}\times P_{jl}}\right)}{n}$$

where q_k is the number of different alleles at locus k, n is the total number of loci in the data set, and pil and pjl are the relative frequencies of allele l in individuals i and j respectively. This measure is identical to that used by (Nei et al. 1983) for population frequency data, but is here applied to pairs of individuals rather than pairs of populations. The measure has properties of taking on values of 0 when genotypes are identical and 1 when genotypes are completely dissimilar.

The spatial genetic structure was calculated using the software Spatial Genetic Software (Degen *et al*. 2001). The spatial genetic structure: Moran's index, Geary's index, number of alleles in common and approaches using genetic distances and F_{ST} values were utilized. The statistical significance of all measures was verified by use of permutation test. All calculated statistics were computed for pairs of data points belonging to a series of spatial distance classes. The Euclidean distance was estimated as a measure of spatial distance between 2 data points. The genetic distances D_N (Nei 1972) or D_G (Gregorius 1978) were utilised for calculation of genetic distograms (Degen and Scholz 1998, Vendramin *et al*. 1999). Genetic distograms represent graphs where mean genetic distances between all pairs of individuals belonging to a spatial distance class (Sq) are plotted against the spatial distance classes. A Monte Carlo permutation procedure was applied to test significant deviations from a spatially random distribution of each calculated measure (Manly 1997). Each permutation consists of a random shuffling of genetic or phenotypic data over the spatial coordinates of the sampled points. For each of the spatial distance classes, observed values are compared with a null distribution, obtained from *N* Monte Carlo trials. Then a user-defined alpha% confidence interval for the parameters was constructed, by ordering the permuted estimates (e.g., Bacilieri *et al*. 1994, Streiff *et al*. 1998).

The demographic features of the Uttar Pradesh (UP) buffalo populations were estimated from the skewness and kurtosis values using the k and g test as given by Reich and Goldstein (1998). The tests are based on distribution and allelic frequency of different alleles at different loci. The tests provide information about the expansion of the population of buffalo in recent past. To test whether the buffalo population of UP is in mutation drift equilibrium, we used the software Bottleneck (Piry *et al.* 1999).

The analysis of molecular variance and assignment of the individuals to the districts of their sampling was carried out using the population assignment test as implemented in the Arlequin software (Excoffier *et al.* 2005). The F_{ST} values were estimated among all the districts of UP and were utilised for multidimensional scaling (MDS) using commercial NTSYSpc software version 2.2 (Exerter software).

RESULTS AND DISCUSSION

The overall population genetic parameters were estimated for 34 district conglomerates for which the data was generated. The allele frequency and the allelic pattern were also estimated (Fig. 2).

The average number of alleles across all the districts of UP was 13.45 for 11 microsatellite loci for which the data was generated. There was large number of allele with very small frequencies in the population while on an average 4.545 alleles had a frequency of more than 5% which resulted in the effective number of alleles to be only 3.878. The average

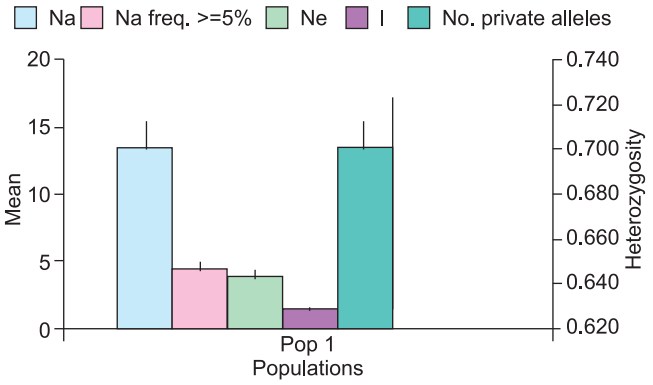

Fig. 2. Graphical representation of allelic patterns across populations

number of alleles in UP buffaloes was much higher compared to other breeds of buffalo in India. The values recently reported are 5.75 alleles in Banni buffaloes (Mishra *et al*. 2009), 4.48 in Marathwada buffaloes (Kathiravan *et al*. 2009), 4.68 in Chilika buffaloes (Mishra *et al*. 2009), and 6.33 in Murrah buffaloes (Bhuyan *et al*. 2010). The higher number of alleles can be attributed to a very large area sampled in this study. The large data set analysed of 625 animals is the major reason as most of the other studies the number of animals on which the data was generated was in the range of 45–50 animals per breed. The effective number of alleles was comparative (3.44 in Banni buffaloes). The effective number of alleles was significantly less in Marathwada buffaloes (2.93). The value of the Shannon information index was 1.516 which is a measure of genetic diversity and the value is significant. The district-wise number of allele recorded for each of the 11 microsatellie loci are being depicted in Fig. 3.

The observed heterozygosity values was also high with an average value of 0.676 and unbiased heterozygosity of 0.677 (corrected for different number of observations). The genetic distances were calculated taking districts as a unit and the values were depicted (Table 2). The UPGMA tree

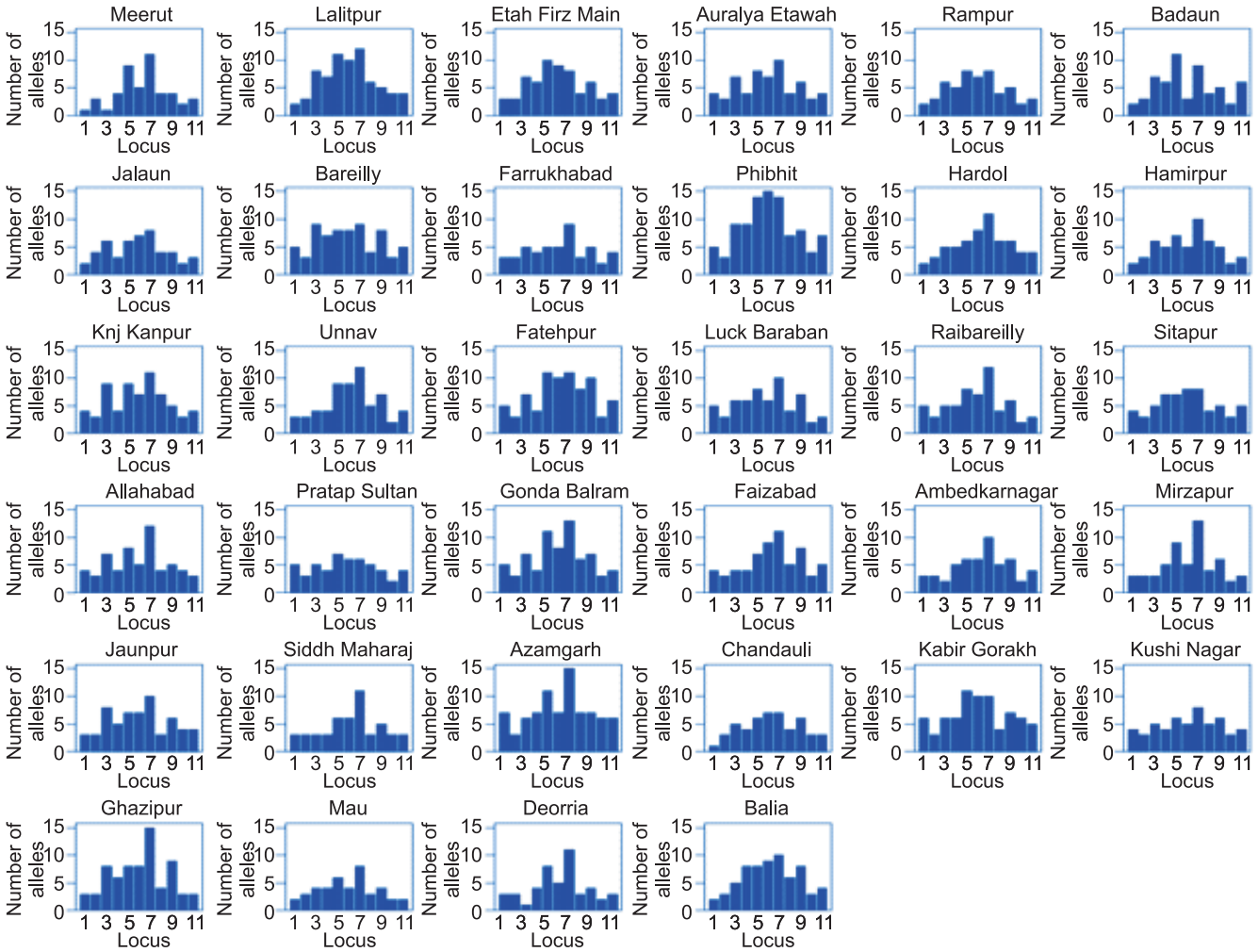

Fig. 3. Allele frequency bar graphs at diffent loci in all the 34 districts

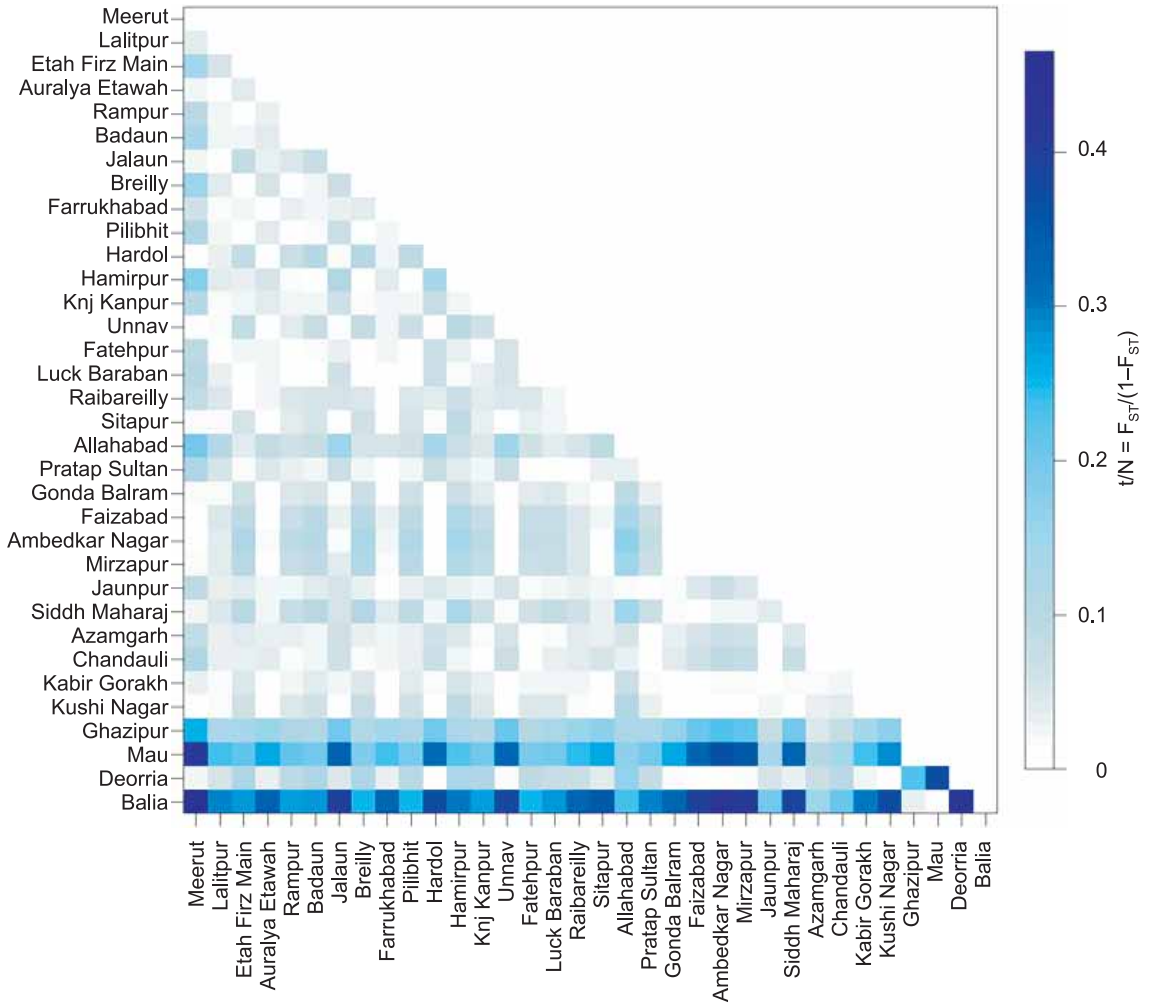

Fig. 4. Slatkin's F_{ST} values

constructed on the basis of Nei's D_A. The tree signifies clustering of the 34 districts of UP into 5 distinctive clusters. The most distinctive cluster is districts of Mau, Balia and Ghazipur (Fig. 5). The other 4 clusters have 3, 6, 8 and 14 districts respectively. An appraisal of location of each of the districts in the geographic map revealed that the districts cluster together with one another even though they do not have contiguity at geographical level. This can be explained by use of semen bulls from different geographical locations under various buffalo improvement programs. However the districts of Mau, Balia and Ghazipur are distinctive, having geographical contiguity and less admixture.

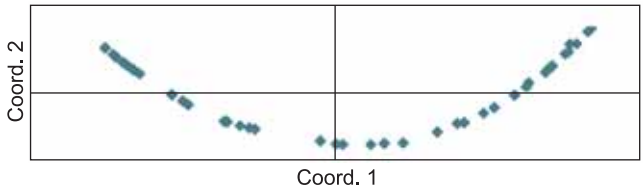

Fig. 5. Principal coordinate analysis showing 5 clusters of population.

The analysis of molecular variance revealed that among districts percentage of variation was 6.62% while among individuals within districts the variation accounted for was 15.80%.

The various districts of UP are continuous in their geography and we intended to find out the most distinctive population structures existing. Even a rough estimate of the *absolute* time since divergence for the various subpopulations if they exist in Indo-Gangetic plains is not known but this parameter may estimate *relative* divergence between a series of pairs of subpopulations. In Fig. 4 the subpopulations and their F_{ST} value are depicted graphically and assuming that the assumptions underlying Slatkin's estimate are fulfilled then *t/N* generations are proportional to the divergence time. The present analysis revealed distinction for Mau and Balia buffaloes from rest of the buffaloes of other districts of Indo-Gangetic plains. The buffaloes of Ghazipur also reveal

Fig. 6. Nei D_A UPGMA with districts taken as population.

consistently large values of F_{ST} with other districts under study. There does not seem to be a distinctive pattern except for these 3 districts.

The principal coordinate analysis also revealed 5 clusters available one of which was a large one and has higher number of districts (Fig. 5). We utilized UPGMA algorithm for the construction of tree using Nei's D_A genetic distance (Fig. 6). The inter-individual genetic distance using Chord distance was utilised for the construction of Neighbor joining tree

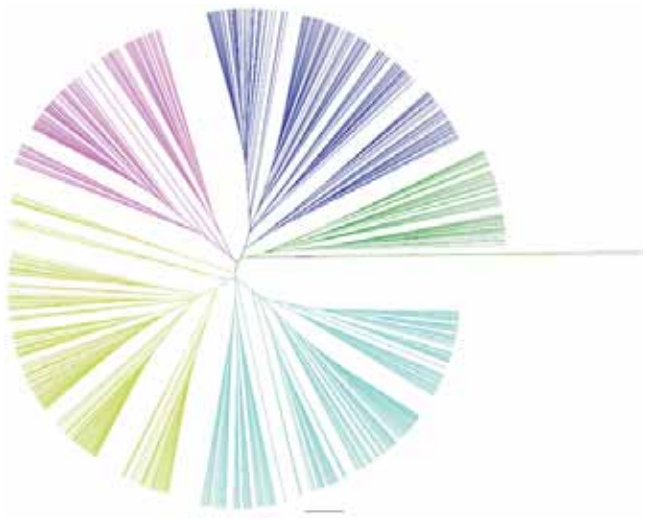

Fig. 7. Inter-individual genetic distances based on Chord distances and Neighbor Joining algorithm.

(Saitou and Nei 1987) and also revealed 5 distinctive clusters (Fig. 7). Thus, the results were very similar for principal coordinate and two genetic distances analysed in the present study.

The multi-dimensional graphs represent that the buffaloes of UP fall in 2 major quadrants (Fig. 8) while at least 4 of the districts are at the extreme ends of the quadrants and are thus distinctive from rest of the districts.

The assignment of buffaloes to their sampling areas was not possible. This was due to the absence of distinctive population structure and lack of distinctiveness in allelic frequencies (Fig.9). Out of 625 buffaloes which were assigned to their sampling areas, only 11% of the buffaloes could be assigned to the areas of their sampling. This shows very small differences in the allelic frequencies in the different districts and absence of a strong population structure.

Demographic parameters of UP buffaloes

The demographic parameters of buffaloes were estimated to find out significant reduction in the effective population size, sudden expansion or recent colonization event. The two attributes were evaluated using all the 3 models of microsatellite evolution and population expansion (Reich and Goldstein 1999). Genetic bottleneck is an evolutionary event in which a significant percentage of a population or species is killed or otherwise prevented from reproducing. Such a situation leads to increase in genetic drift, as the rate of drift is inversely proportional to the population size. In the real genetic terms the alleles in the homozygous conditions shall be reduced which shall lead to heterozygosity excess. The significant heterozygosity excess shall reveal the occurrence of severe reduction in effective population size. The 3 tests, viz. Sign test, Standardised differences test and Wilcoxon Rank test were used to test the significance of heterozygosity excess under the 3 models of microsatellite evolution, Infinite Allele Model (IAM), Two Phase Model (TPM) and Stepwise Mutation Model (SMM). In the first model IAM, the expected number of loci with heterozygosity excess was 6.47 and the number of loci with heterozygosity excess was 8. The population shows, there is significant excess in IAM. Similar results were obtained in standardized differences test and Wilcoxon Rank test with T_2 value of 1.263 and probability of 0.10333 while the probability value of Wilcoxon one tail test was 0.06152. However the microsatellites are not known to evolve indefinitely and thus the test results are too liberal in detecting heterozygosity excess. The Two Phase Model resulted in 10 loci with heterozygosity deficiency and T_2 values as –3.734 and Wilcoxon Rank test probability of 0.00342 which is significant but in favor of heterozygotic deficiency. The values for SMM which is most conservative model of microsatellite evolution revealed 10 loci with heterozygotic deficiency and T_2 value of –14.098 and Wilcoxon rank test probability value of 0.0. Thus the tests

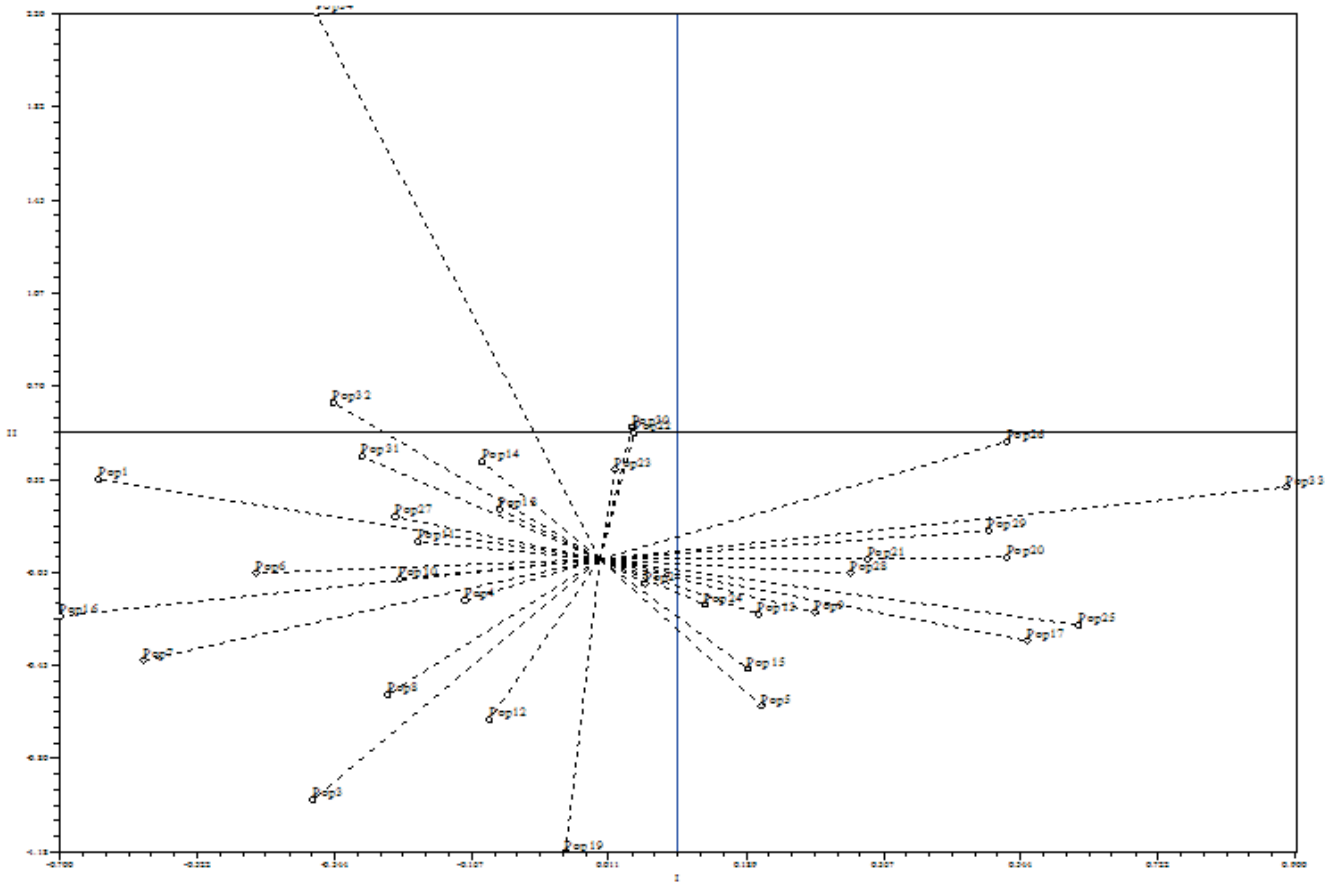

Fig. 8. Multidimensional scaling graph using F_{ST} values. The various districts have been numbered as pop1–34. The pop description is given in Table 1.

revealed there is no heterozygosity excess rather a heterozygosity deficiency. The tests point towards no recent colonisation event or recent reduction in the effective population size. The mode shift test also reveals the same. The mode shift test reveals a normal L shaped curve as depicted below in Fig.1.

The *k*-test of Reich and Goldstein (1998) exploits differences between the expected distributions of alleles in populations at Mutation Drift Equilibrium and populations that have recently expanded. The *g*-test of Reich and Goldstein (1998) which compares the between-loci variance in the number of repeats with a theoretical expectation derived assuming that the loci follow SMM and that the population size is stable. We performed both the *k*- and the *g*-tests. *k*-statistics were calculated for each locus, and the significance of the proportion of positive *k* values was based on a binomial distribution with the probability of a positive *k* set conservatively as 0.515 (Reich *et al.* 1999). Significance levels for the *g*-test were compared to the values given in Reich *et al.* 1999.

The inter-locus test (g test) was conducted and the estimated value was found to be 1.226735. This was higher than the fifth percentile cut off of g value (Reich *et al*. 1999). Thus the inter locus g test reveals the constancy of population size (Table 2).

The within locus test (k test) showed 6 loci to be with negative values while 5 loci had positive values. The number of loci with negative value were thus greater than the number of loci with the positive value and thus the null hypothesis of constancy of population size of UP buffaloes was accepted (P= 0.45946).

Isolation by distance: The Mantel test was carried out using the software Isolation by Distance (IBD) (Bohonak 2002). The test was carried out to assess Isolation by distance whether more distant population pairs are more different genetically, importance of specific barriers to gene flow, effects of population history from ongoing gene flow, evaluation of explanatory power of alternative dispersal pathways can be tested. The significance in the isolation by distance relationship was tested by generating a null distribution by randomizing rows and columns of one matrix while holding the other constant. The Mantel test provides only an assessment whether the association is significant. The Regression techniques was used to estimate the slope and intercept of the IBD relationship. Reduced major axis (RMA) regression was utilised in the present study as

Table 1. Table depicts the Nei's genetic distance (below diagonal) and F_{ST} above diagonal

	Pop1	Pop2	Pop3	Pop4	Pop5	Pop6	Pop7	Pop8	Pop9	Pop10	Pop11	Pop12	Pop13	Pop14	Pop15	Pop16	Pop17
Pop1	0.000	0.042	0.122	0.027	0.089	0.116	0.022	0.130	0.057	0.109	0.001	0.149	0.095	0.003	0.085	0.094	0.075
Pop2	0.290	0.000	0.049	0.008	0.026	0.028	0.014	0.036	0.018	0.028	0.035	0.038	0.016	0.019	0.013	0.031	0.043
Pop3	0.197	0.099	0.000	0.041	0.014	0.026	0.080	0.009	0.024	0.009	0.077	0.029	0.027	0.076	0.025	0.004	0.013
Pop4	0.198	0.142	0.131	0.000	0.029	0.038	0.031	0.051	–0.003	0.036	0.013	0.051	0.036	0.012	0.026	0.023	0.028
Pop5	0.239	0.236	0.208	0.172	0.000	0.014	0.043	0.000	0.029	0.004	0.064	0.017	0.027	0.038	0.008	–0.004	0.048
Pop6	0.331	0.399	0.347	0.172	0.313	0.000	0.072	0.022	0.024	0.015	0.096	0.003	0.024	0.073	0.015	0.011	0.049
Pop7	0.170	0.165	0.117	0.120	0.156	0.268	0.000	0.067	0.030	0.062	0.015	0.102	0.056	0.007	0.033	0.058	0.050
Pop8	0.242	0.214	0.198	0.223	0.242	0.330	0.194	0.000	0.036	0.000	0.094	0.010	0.021	0.078	0.011	0.005	0.047
Pop9	0.336	0.067	0.125	0.188	0.273	0.427	0.208	0.288	0.000	0.026	0.022	0.042	0.027	0.024	0.022	0.006	–0.010
Pop10	0.174	0.172	0.139	0.146	0.164	0.308	0.090	0.224	0.224	0.000	0.084	0.008	0.023	0.067	0.007	–0.003	0.048
Pop11	0.249	0.065	0.102	0.136	0.256	0.367	0.170	0.201	0.109	0.162	0.000	0.119	0.073	0.006	0.065	0.067	0.036
Pop12	0.241	0.129	0.096	0.193	0.210	0.388	0.140	0.245	0.152	0.146	0.166	0.000	0.028	0.092	0.031	0.020	0.065
Pop13	0.228	0.153	0.130	0.123	0.185	0.292	0.116	0.197	0.199	0.137	0.158	0.159	0.000	0.056	0.002	0.034	0.035
Pop14	0.278	0.299	0.248	0.121	0.232	0.078	0.189	0.261	0.336	0.263	0.285	0.284	0.210	0.000	0.052	0.054	0.044
Pop15	0.186	0.069	0.088	0.111	0.221	0.350	0.132	0.201	0.105	0.140	0.078	0.136	0.122	0.265	0.000	0.008	0.043
Pop16	0.213	0.200	0.164	0.187	0.183	0.320	0.127	0.253	0.226	0.161	0.218	0.183	0.154	0.231	0.168	0.000	0.023
Pop17	0.263	0.058	0.102	0.145	0.250	0.353	0.162	0.204	0.098	0.140	0.071	0.125	0.162	0.275	0.090	0.234	0.000
Pop18	0.297	0.081	0.091	0.187	0.225	0.397	0.154	0.201	0.111	0.174	0.127	0.113	0.140	0.286	0.125	0.193	0.097
Pop19	0.206	0.205	0.183	0.146	0.206	0.255	0.181	0.225	0.229	0.217	0.194	0.188	0.179	0.203	0.162	0.230	0.197
Pop20	0.190	0.078	0.088	0.108	0.178	0.323	0.127	0.172	0.125	0.121	0.077	0.126	0.111	0.234	0.063	0.158	0.095
Pop21	0.162	0.128	0.121	0.124	0.184	0.324	0.103	0.198	0.203	0.123	0.143	0.143	0.118	0.232	0.102	0.127	0.145
Pop22	0.244	0.057	0.116	0.142	0.201	0.378	0.162	0.220	0.088	0.166	0.098	0.157	0.166	0.276	0.087	0.189	0.094
Pop23	0.238	0.096	0.084	0.124	0.167	0.290	0.112	0.210	0.126	0.135	0.126	0.111	0.106	0.191	0.099	0.145	0.097
Pop24	0.183	0.151	0.129	0.126	0.161	0.283	0.092	0.229	0.177	0.085	0.167	0.145	0.127	0.226	0.129	0.157	0.136
Pop25	0.345	0.438	0.388	0.210	0.321	0.070	0.314	0.361	0.473	0.360	0.412	0.402	0.349	0.076	0.395	0.359	0.404
Pop26	0.310	0.037	0.103	0.159	0.253	0.413	0.194	0.249	0.067	0.191	0.077	0.156	0.172	0.311	0.084	0.218	0.061
Pop27	0.273	0.053	0.087	0.141	0.218	0.393	0.167	0.237	0.076	0.168	0.086	0.105	0.135	0.279	0.075	0.184	0.088
Pop28	0.181	0.168	0.115	0.122	0.157	0.275	0.064	0.216	0.187	0.078	0.164	0.149	0.098	0.212	0.120	0.118	0.147
Pop29	0.233	0.147	0.172	0.137	0.180	0.364	0.154	0.217	0.204	0.168	0.177	0.175	0.157	0.261	0.125	0.185	0.190
Pop30	0.183	0.104	0.114	0.130	0.198	0.341	0.112	0.221	0.156	0.085	0.120	0.101	0.159	0.246	0.103	0.193	0.095
Pop31	0.194	0.136	0.118	0.148	0.153	0.301	0.091	0.200	0.171	0.101	0.163	0.131	0.140	0.219	0.129	0.136	0.144
Pop32	0.295	0.077	0.124	0.167	0.257	0.404	0.200	0.251	0.097	0.197	0.093	0.160	0.172	0.305	0.099	0.236	0.096
Pop33	0.226	0.068	0.097	0.126	0.196	0.380	0.120	0.224	0.121	0.116	0.104	0.106	0.144	0.275	0.092	0.203	0.086
Pop34	0.253	0.028	0.083	0.130	0.217	0.362	0.132	0.207	0.063	0.142	0.067	0.114	0.131	0.263	0.064	0.167	0.061

Pop1	Meerut	Pop5	Rampur	Pop9	Farrukhabad	Pop13	KnjKanpur
Pop2	Lalitpur	Pop6	Badaun	Pop10	Pilibhit	Pop14	Unnav
Pop3	EtahFirzMain	Pop7	Jalaun	Pop11	Hardoi	Pop15	Fatehpur
Pop4	AuraiyaEtawah	Pop8	Bareilly	Pop12	Hamirpur	Pop16	LuckBaraban
						Pop17	Raibareily

implemented in the IBD software.

The partial correlations between genetic patterns, geographic distance and a third variable matrix (Indicator variable) was done to find out if the presence of flowing rivers in the Indo-Gangetic plains has significant relationship with genetic distances or the major rivers flowing in the region result in genetic structure observed in analysis. This technically reveals if a particular environmental gradient correlate better with gene flow than direct-line geographical distance.

The Mantel test of matrix correlation between the genetic distances and the geographic distances revealed a correlation value (r) of 0.1781 and the p value associated with this based on 1000 randomization were 0.0240 implying significant relationship between the genetic and geographical distance. The correlation of genetic and indicator matrix was found to be r= 0.0809 with a p value of 0.1470 showing the correlation to be nonsignificant. The partial correlation of genetics and geographical distance controlling the indicator matrix was r =0.1738 with a p value of 0.0250 which is significant at 5% level of significance. The partial correlation of genetics and indicator matrix controlling geography was non significant ($P \leq 0.1750$). Similar trend was observed when Mantel test for matrix correlation between genetic distance and log (geographical distance) was estimated. The reduced major axis regression revealed a slope of 0.000694 and an R^2 value

(Concluded Table 1)

	Pop18	Pop19	Pop20	Pop21	Pop22	Pop23	Pop24	Pop25	Pop26	Pop27	Pop28	Pop29	Pop30	Pop31	Pop32	Pop33	Pop34
Pop1	0.021	0.168	0.107	0.020	0.007	–0.004	0.016	0.083	0.022	0.079	0.109	0.033	0.005	0.205	0.294	0.021	0.317
Pop2	0.017	0.098	0.049	0.020	0.042	0.037	0.035	0.034	0.043	0.029	0.034	0.019	0.015	0.115	0.192	0.053	0.226
Pop3	0.049	0.041	0.017	0.058	0.083	0.107	0.090	0.039	0.091	0.036	0.030	0.048	0.058	0.120	0.176	0.104	0.220
Pop4	0.003	0.080	0.043	0.005	0.013	0.020	0.005	0.025	0.023	0.030	0.040	0.002	0.008	0.137	0.211	0.029	0.252
Pop5	0.037	0.061	0.029	0.048	0.062	0.081	0.073	0.023	0.069	0.030	0.015	0.025	0.043	0.109	0.174	0.082	0.215
Pop6	0.055	0.070	0.024	0.053	0.087	0.096	0.085	0.040	0.090	0.026	0.023	0.039	0.059	0.104	0.168	0.108	0.221
Pop7	0.008	0.128	0.064	0.020	0.034	0.014	0.036	0.052	0.054	0.058	0.061	0.023	–0.003	0.166	0.254	0.043	0.289
Pop8	0.058	0.050	0.027	0.067	0.090	0.115	0.103	0.031	0.103	0.031	0.015	0.048	0.069	0.100	0.158	0.109	0.201
Pop9	–0.006	0.050	0.004	0.004	0.030	0.025	0.014	–0.004	0.035	0.021	0.028	–0.008	0.004	0.124	0.191	0.029	0.248
Pop10	0.052	0.056	0.024	0.057	0.085	0.099	0.088	0.032	0.086	0.032	0.032	0.039	0.060	0.105	0.164	0.088	0.198
Pop11	0.015	0.119	0.072	0.011	–0.004	–0.002	0.012	0.042	0.023	0.058	0.062	0.019	–0.004	0.165	0.239	0.014	0.272
Pop12	0.086	0.062	0.036	0.064	0.104	0.125	0.103	0.034	0.115	0.042	0.023	0.053	0.081	0.112	0.187	0.115	0.231
Pop13	0.033	0.048	0.022	0.040	0.071	0.084	0.078	0.026	0.055	0.017	0.007	0.029	0.046	0.097	0.165	0.107	0.214
Pop14	0.006	0.128	0.066	0.010	0.010	0.001	0.013	0.049	0.020	0.054	0.065	0.013	–0.009	0.169	0.244	0.016	0.275
Pop15	0.030	0.058	0.010	0.036	0.068	0.076	0.067	0.020	0.057	0.012	0.009	0.024	0.043	0.096	0.159	0.086	0.199
Pop16	0.024	0.041	–0.003	0.043	0.069	0.080	0.070	0.024	0.076	0.020	0.031	0.022	0.043	0.112	0.168	0.072	0.219
Pop17	0.014	0.049	0.012	0.023	0.048	0.048	0.047	0.029	0.058	0.035	0.040	0.022	0.016	0.135	0.195	0.065	0.247
Pop18	0.000	0.083	0.030	0.010	0.027	0.009	0.011	0.028	0.030	0.032	0.050	0.002	–0.007	0.144	0.209	0.041	0.261
Pop19	0.212	0.000	0.032	0.081	0.119	0.146	0.131	0.020	0.128	0.051	0.033	0.067	0.099	0.112	0.143	0.143	0.191
Pop20	0.125	0.154	0.000	0.030	0.067	0.079	0.062	0.014	0.063	0.010	0.012	0.021	0.033	0.113	0.168	0.077	0.225
Pop21	0.138	0.197	0.105	0.000	0.004	0.005	0.007	0.020	0.015	0.028	0.037	0.000	–0.005	0.143	0.209	0.021	0.245
Pop22	0.098	0.195	0.097	0.133	0.000	0.003	0.008	0.042	0.013	0.054	0.058	0.013	0.001	0.173	0.247	0.010	0.284
Pop23	0.076	0.184	0.100	0.088	0.105	0.000	0.002	0.064	0.023	0.066	0.086	0.017	–0.010	0.189	0.270	0.013	0.300
Pop24	0.160	0.202	0.116	0.140	0.147	0.121	0.000	0.044	0.022	0.058	0.079	0.018	–0.002	0.179	0.260	0.017	0.291
Pop25	0.433	0.268	0.367	0.375	0.403	0.336	0.333	0.000	0.041	0.005	–0.012	0.005	0.027	0.071	0.120	0.050	0.167
Pop26	0.088	0.218	0.102	0.156	0.075	0.096	0.169	0.458	0.000	0.046	0.069	0.016	0.014	0.165	0.249	0.033	0.281
Pop27	0.110	0.190	0.083	0.146	0.076	0.097	0.147	0.421	0.057	0.000	0.012	0.021	0.035	0.047	0.093	0.065	0.130
Pop28	0.153	0.190	0.111	0.102	0.149	0.091	0.072	0.344	0.183	0.152	0.000	0.022	0.042	0.068	0.124	0.091	0.173
Pop29	0.183	0.223	0.132	0.141	0.130	0.155	0.122	0.377	0.182	0.136	0.137	0.000	0.005	0.125	0.190	0.028	0.230
Pop30	0.136	0.177	0.094	0.119	0.113	0.119	0.096	0.357	0.125	0.109	0.125	0.132	0.000	0.148	0.225	–0.003	0.272
Pop31	0.126	0.204	0.114	0.122	0.130	0.122	0.088	0.338	0.160	0.143	0.098	0.150	0.119	0.000	–0.004	0.188	0.032
Pop32	0.123	0.189	0.112	0.165	0.091	0.133	0.184	0.435	0.078	0.093	0.193	0.188	0.136	0.179	0.000	0.273	–0.011
Pop33	0.105	0.169	0.081	0.108	0.092	0.111	0.106	0.406	0.092	0.085	0.125	0.142	0.072	0.108	0.115	0.000	0.301
Pop34	0.073	0.185	0.067	0.102	0.054	0.076	0.116	0.405	0.041	0.052	0.121	0.138	0.093	0.099	0.072	0.061	0.000

Pop18	Sitapur	Pop22	Faizabad	Pop26	SiddhMaharj	Pop30	KushiNgr
Pop19	Allahabad	Pop23	AmbdkarNgr	Pop27	Azamgarh	Pop31	Ghazipur
Pop20	PratapSultan	Pop24	Mirzapur	Pop28	Chandauli	Pop32	Mau
Pop21	GondaBalram	Pop25	Jaunpur	Pop29	KabirGorakh	Pop33	Deorria
						Pop34	Balia

of 0.0317 which showed that only a small percent of variation is explained by the regression. When the data of geographical and genetic distance was plotted it almost revealed a parallel line along the X axis.

Similar analysis was carried out using Mantel test (Fig. 11) for matrix correlation between genetic distance and log of geographical distance. The correlation of genetic and geographical distance was found to be 0.1593 with p value of 0.0190 from 1000 randomizations, thus the correlation was found to be significant and the correlation of genetic and indicator matrix nonsignificant (r =0.0809 with P≤0.1500). The partial correlation of genetic and geographical distance controlling the indicator matrix was 0.1578 (P≤0.0190) which was significant. The partial correlation of genetics and indicator matrix controlling geography was nonsignificant (r =0.0778, P≤0.1610) The RMA regression analysis gave a slope of 0.3646 with R^2 value of 0.0254. The study revealed that the differences among the individuals increase with increasing geographic distances.

Allelic aggregation index: The allelic aggregation index was estimated which is based on the statistical concept of aggregation. Aggregation indices are commonly used in ecological studies to characterize spatial distributions of individuals across landscapes (Clark and Evans 1954, Hopkins and Skellam 1954, Pielou 1977) and have been

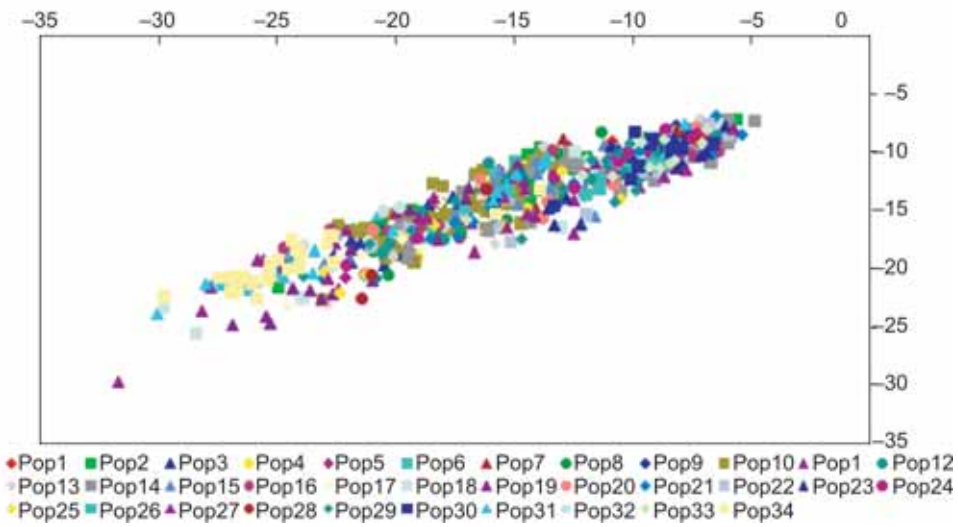

Fig. 9. Population assignment. Populations are defined in Table 1.

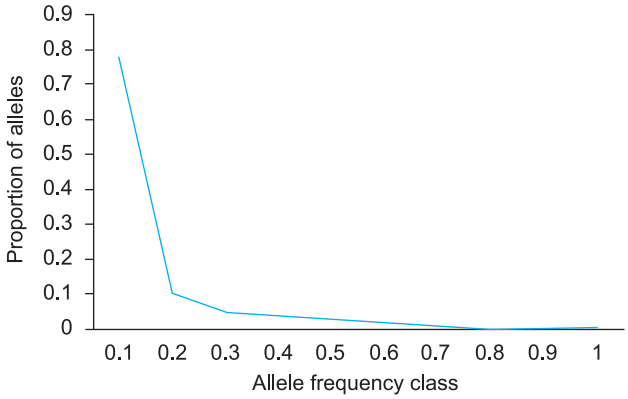

Fig. 10. Normal L shaped curve indicating no bottleneck.

widely used specifically with respect to describing the presence of either randomized, clumped, or uniform spatial distributions of individuals. Allelic aggregation index provides a basis for testing the null hypothesis that each allele at a locus is distributed at random across a landscape (i.e. no aggregation or genetic structure) relative to the aggregation of the actual animals. In the present analysis the spatial aggregation R values were 0.02658 with global aggregation of allele and loci (Rave) as 0.70263. The values were statistically highly significant (P<0.0). It means the alleles at the 11 loci in buffaloes of Uttar Pradesh are not randomly distributed along the landscape but forms aggregations which point towards existence of population structures.

Genetic landscape of UP buffaloes: Alleles *In Space* also implements a novel technique that can be used to obtain graphical representations of genetic distance patterns across landscapes. The three-dimensional surface plots generated by this procedure are referred to as ''genetic landscape shapes''. The three-dimensional surface plots where X and Y coordinates correspond to geographical locations on the rectangular grid and surface plot heights (Z) reflect genetic distances. The genetic landscape of the Indo-Gangetic plain is shown as Fig. 12. Basically, it contains an inferred graphical

Table 2. Showing interiloucs (g test) and intralocus test values

Locus	K	n	Mean	Variance	sum2	sum4	sig4	gamma4
BMS2684	–90.617	1250	95.688	3.050697	3810.32	142756	9.222494	114.5267
BMS2722	2.590998	1250	110.9808	2.030055	2535.539	10330.55	4.117797	8.27115
BMS2785	4884.187	1250	103.5296	66.35341	82875.4	7674134	4401.38	6137.841
CSSM08	197.9142	1250	188.4272	20.36019	25429.88	1053198	414.1935	843.2696
CSSM19	–1395.29	1250	138.736	47.29934	59076.88	8725535	2233.418	6992.077
CSSM43	19221.15	1250	242.5376	182.1991	227566.7	79643777	33172.02	63759.9
CSSM47	11372.05	1250	140.8224	151.0493	188660.6	57034281	22797.57	45664.17
ILSTS05	–6308.89	1250	180.568	36.69553	45832.72	12055533	1339.897	9668.911
ILSTS11	–5041.73	1250	265.028	15.29865	19108.02	7004357	229.7361	5620.349
ILSTS49	–10646.6	1250	136.4808	6.719007	8392.039	13382203	36.58234	10739.91
RM232	–108.321	1250	115.1168	3.563609	4450.947	175525.2	12.59672	140.8099

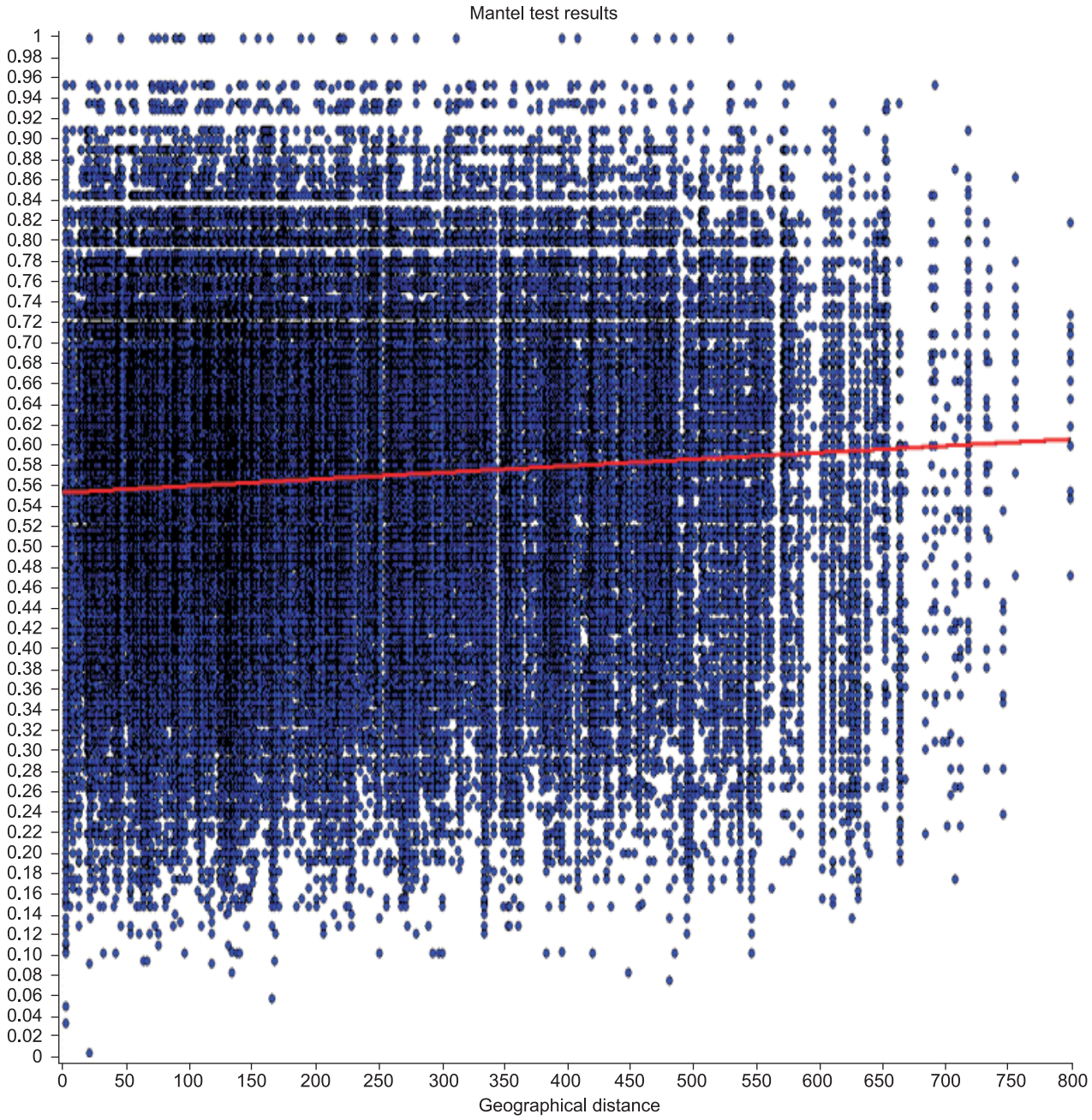

Fig. 11. Matrix correlations between geographical distance and genetic distance (Mantel test).

representation of patterns of diversity across the sampled landscape, that ideally contains peaks in areas where there are large genetic distances.

Genetic landscape shapes portray patterns of genetic diversity/divergence across landscapes, with a goal to identify spatial patterns associated with the largest genetic distances in a data set. Since a correlation between genetic and geographical distances existed for the data set, it was thought appropriate to account for this correlation. We utilised the calculation of "pseudoslopes" from the genetic and geographical distance matrix. These "pseudoslopes" were derived by the AIS program as the quotient of congruent elements from the genetic and geographical distance matrices.

Ten distance classes made are shown in the column 1 (Table 3, Fig. 13). The values of the Moran index are depicted in column 2. The negative values of Moran's index give negative spatial auto-correlation while the positive values give positive auto-correlation. The positive values signify that for the first 2 distance classes and the sixth distance

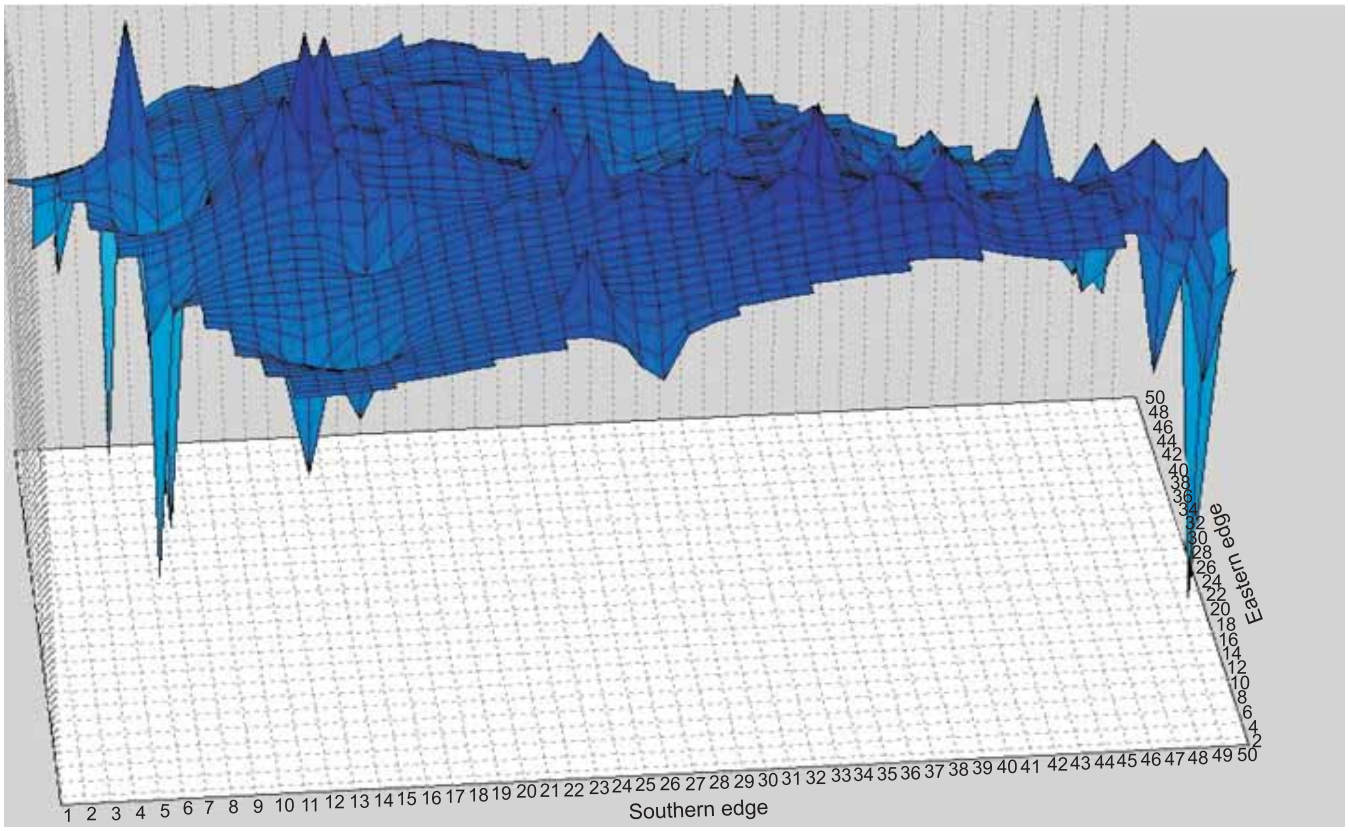

Fig. 12. Genetic landscape of Indo-Gangetic plain.

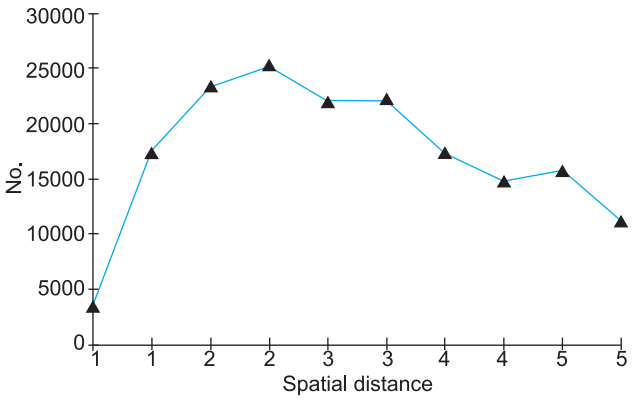

Fig. 13. Pairs of data per distance class.

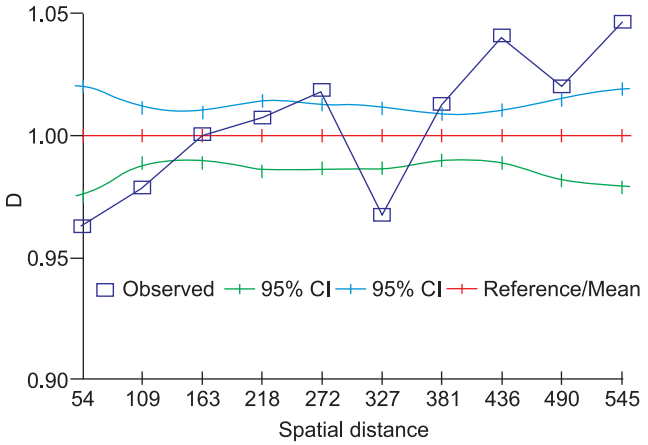

Fig. 14. Correlogram using Moran's index.

class (272.45 to 326.94) there is positive auto correlation while rests of the values have negative autocorrelation. The correlogram showing the Moran Index is depicted in Fig. 14.

Similarly the Geary's C values lie between 0 and 2 (Fig. 15). The value of 1 shows no spatial autocorrelation while the values of less than 1 mean positive autocorrelation and that larger than 1 mean negative autocorrelation. There are three distance classes which showed values lesser than 1 showing a positive auto-correlation. The results are very similar for both the Geary' C and the Moran index. Geary's C is inversely related to Moran's I, but it is not identical. Moran's I is a measure of global spatial autocorrelation, while Geary's C is more sensitive to local spatial autocorrelation. The two tests revealed that there is deviation from the randomness at spatial level which point towards the existence of population structure.

Table 3. Distance classes, observed D for Moran and Geary's index

Distance	D(obs.) Moran index	P(D)< (-CI)	D(obs.) Geary's index	P(D)< (-CI)
0–54.49	0.051142	0.00	0.963282	1.00
54.49–108.98	0.017196	0.00	0.978809	1.00
108.98–163.47	–0.005945	1.00	1.000187	0.518
163.47–217.96	–0.015451	1.00	1.007582	0.176
217.96–272.45	–0.003936	0.952	1.018088	0.004
272.45–326.94	0.001934	0.010	0.967556	1.00
326.94–381.43	–0.017282	1.00	1.012647	0.004
381.43–435.92	–0.035318	1.00	1.040577	0.000
435.92–490.41	–0.017979	1.00	1.019632	0.000
490.41–544.9	–0.011108	1.00	1.046536	0.000

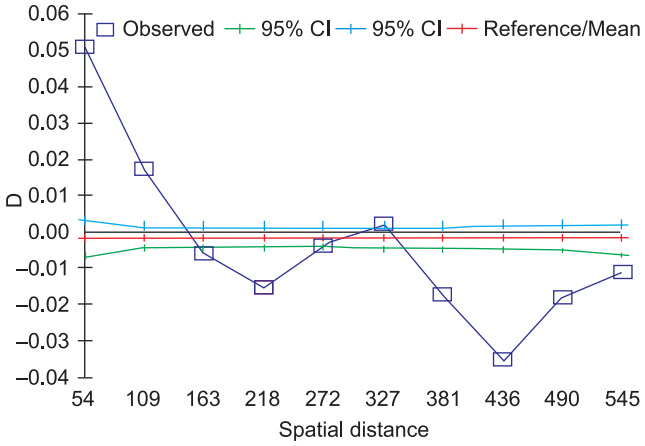

Fig. 15. Correlogram using Geary's index.

The study revealed that there is significant correlation between the geographic and genetic distances as revealed by Mantel test. The partial Mantel test revealed no significant effect of existence of environmental indicator variable (rivers and their tributaries in this present study) with the genetic distances meaning the rivers do not contribute to the non-randomness of the allelic patterns obtained in the study. There is however significant allelic aggregation in the buffalo populations with significant spatial patterns as evidenced by Moran's and Geary's index. Significant genetic structure has been evidenced only for 3 districts of Mau, Balia and Ghazipur. Thus there are local spatial genetic population structuring evidenced by Moran and Geary's Index and detailed analysis of spatial genetic structures using Wombling, genetic barrier studies and genetic bandwidth mapping are advocated.

REFERENCES

Bacilieri R, Labbe T and Kremer A. 1994. Intraspecific genetic structure in a mixed population of *Quercus petraea* (Matt.) Liebl. and *Q. robur* L. *Heredity* **73**: 130–41.

Bhuyan D K, Sangwan M L, Gole V C and Sethi R K. 2010. Studies on DNA fingerprinting in Murrah buffaloes using microsatellite markers. *Indian Journal of Biotechnology* **9**: 367–70.

Bohonak A J. 2002. IBD (Isolation by Distance): A program for analyses of isolation by distance. *Journal of Heredity* **93**: 153–54.

Brouns G, De Wulf A and Constales D. 2003. Delaunay triangulation algorithms useful for multibeam echosounding. *Journal of Surveying Enginnering* **129**: 79–84.

Clark P J and Evans F C. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* **35**: 445–53.

Degen B and Scholz F. 1998. Spatial genetic differentiation among populations of European beech (*Fagus sylvatica* L.) in Weste rn Germany as identified by geostatistical analysis. *Forest Genetics* **5**: 191–99.

Degen B. 2000. SGS: Spatial Genetic Software. Computer program http://kourou.cirad.fr/genetique/software.html.

Excoffier L, Laval G and Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**:47–50.

Gregorius H R. 1978. The concept of genetic diversity and its formal relationship to heterozygosity and genetic distance. *Mathematical Bioscience* **41**: 253–71.

Hopkins B and Skellam J G. 1954. A new method for determining the type of distribution of plant individuals. *Annals of Botany* **18**: 213–27.

Kathiravan P, Mishra B P, Kataria R S and Sadana D K. 2009. Evaluation of genetic architecture and mutation drift equilibrium of Marathwada buffalo population in central India. *Livestock Science* **121**: 288–93.

Manly B F J. 1997. Randomization, Bootstrap and Monte Carlo Methods in Biology. 2nd edn. Chapman and Hall London.

Miller M P. 2005. Alleles in space (AIS): Computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity* **96**: 722–24.

Mishra B P, Kataria R S, Kathiravan P, Bulandi S S, Singh K P and Sadana D K. 2009. Evaluation of genetic variability and mutation drift equilibrium of Banni buffalo using multi locus microsatellite markers. *Tropical Animal Health Production* **41**: 1203–11.

Mishra B P, Kataria R S, Bulandi S S, Prakash B, Kathiravan P, Mukesh M and Sadana D K. 2009. Riverine status and genetic structure of Chilika buffalo of Eastern India as inferred from cytogentic and molecular marker-based analysis. *Journal of Animal Breeding and Genetics* **126**: 69–79.

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

Nei M. 1972. Genetic distance between populations. *American Naturalist* **106**: 283–92.

NTSys PC version 2.2 (Exerter software) 47 Route 25A, Suite 2, Setauket, NY 11733–2870 USA.

Pielou E C. 1977. *Mathematical Ecology* Pp.385. Wiley, New York.

Piry S, Luikart G and Cornuet J-M. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**: 502–03.

Reich D E and Goldstein D B. 2001. Detecting Association in a

Case-Control Study While Correcting for Population Stratification, *Genetic Epidemiology* **20**: 4–16.

Reich D E, Feldman M W and Goldstein D B. 1999. Statistical properties of two tests that use multilocus data sets to detect population expansions. *Journal of Molecular Biological Evolution* **16**: 453–466.

Reich D E and Goldstein D B. 1998. Genetic evidence for a Paleolithic human population expansion in Africa. *Proceedings of National Academic Science USA.* **95**: 8119–23.

Saitou N and Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Journal of Molecular Biological Evolution* **4**: 406–25.

Sambrook J, Fritsch E F and Maniatis T. 1989. Molecular Cloning: A Laboratory Manual.-Cold Spring Laboratory Press, Cold Spring Harbour-New York.

Slatkin M and Voelm L. 1991. FST in a hierarchical island model. *Genetics* **127**: 627–29.

Streiff R, Labbe T, Bacilieri R, Steinkellner H, Glossl J and Kremer A. 1998. Within-population genetic structure in Quercus robur L. and Quercus petraea (Matt.) Liebl assessed with isozymes and microsatellites. *Molecular Ecology* **7**: 317–28.

Vendramin G G, Degen B, Petit R J, Anzidei M, Madaghiele A and Ziegenhagen B. 1999. High level of variation at *Abies alba* chloroplast microsatellite loci in Europe. *Molecular Ecology* **8**:1117–26.

Watson D F and Philips G M. 1985. A refinement of inverse distance weighted interpolation. *Geo-Processing* **2**: 315–27.

Watson D F. 1992. *Contouring: A guide to The Analysis and Display of Spatial data.* Pergamon Press, ISBN No. 0080402860 New York.