Short Communications

Indian Journal of Animal Sciences **82** (11): 1434–1437, November 2012

Microsatellite analysis of buffaloes of Indo-Gangetic Plains

JYOTI JOSHI[1], R K SALAR[2], P BANERJEE[3], S B GOKHALE[4], UPASNA S[5], UMA GAUR[6], M S TANTIA[7] and R K VIJH[8]

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

Received: 29 April 2011; Accepted: 2 May 2012

Key words: Buffalo, Microsatellite, Genetic distance, Phylogenetic tree

Buffalo- a member of family Bovidae is a major contributor to food security of poor farmers for Indian sub-continent and South-East Asia. It not only contributes to milk, meat and draught power but is also a major source of skin and hides.

Buffaloes are more economical than cattle under low mechanization. India has 105.4 million buffaloes (Anonymous 2003) with 10 distinct populations/breeds viz: Bhadawari, Jaffarabadi, Marathwada, Mehsana, Murrah, Nagpuri, Nili-Ravi, Pandharpuri, Surti and Toda (Sahai and Vijh 2000). There are also several other distinctive populations not covered under the definition of breeds.

Buffalo accounts for the 50% of total milk production (52.07 mt) and nearly 90% of meat production in 2005 – 2006 (DAH 2006). Out of a total population of 105.4 million heads of buffaloes in India, Uttar Pradesh has largest number of buffalo heads which 23 million constituting 26.1% to the total buffalo population of the country. Yet there is only one defined breed of buffalo i.e; Bhadawari and rest other buffaloes are considered as Murrah type/grade undefined populations (Sahai and Vijh 2000). Authors have made an attempt to characterize the buffalo germplasm of Indo-Gangetic plains, and evaluate population structure if it exists.

Microsatellites have been markers of choice to identify the population structures because of their high polymorphic information content and have been utilized in the present study.

Blood samples of 625 buffaloes were collected from the 34 districts of Uttar Pradesh. This was done to ensure that the complete landscape of Ganga, Yamuna, Chambal and *Tarai* region is covered. The animals were selected at random and 10 ml of whole blood was collected from the jugular vein in EDTA-coated Vacutainer tubes and transported to the laboratory at 0° to 5°C.

Present address: [1,3,6]Senior Research Fellow, [5]Research Associate, [7,8]Principal Scientist, NBAGR, Karnal, 132001. [2]Reader, Chaudhary Devi Lal University, Sirsa-125055 Haryana.

[4]Vice President, BAIF, Pune.

DNA was extracted from whole blood by using a standard protocol (Roche diagnostics). The concentration of DNA was judged by comparison with the standard DNA marker concentration on agarose gels. The quality of DNA was checked on 0.6% agarose gels prepared in TRIS-acetate EDTA buffer.

A total of 11 heterologous micro-satellite loci were chosen for the study. These loci were ILSTS011, CSSM43, ILSTS05, ILSTS049, CSSM47, CSSM08, CSSM19, BMS2722, BMS2684, BMS2785 and RM232. The criterion for selection of the heterologous microsatellite loci was based on their polymorphism in buffaloes, polymorphism information content value and number of alleles (Navani *et al*. 2002). The 5' ends of the forward primers were labeled with either VIC, NED, FAM or PET (Applied Biosystems, Foster City, CA) dyes. The 11 loci were located on different chromosome in cattle.

The PCR conditions were standardized for all of the 11 primer pairs selected for the study. Polymerase chain reaction amplification was carried out in a 20 µl reaction containing 50 ng of genomic DNA, 150-µM dNTP, 5 p mol each of forward (labeled) and reverse primers, 1 U of Taq DNA polymerase, and 1x reaction buffer (containing 1.5-mM $MgCl_2$). Amplification was carried out by using an ABI 9700 instrument (Applied Biosystems) with initial denaturation at 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 57 to 60°C (primer specific) for 45 s and extension at 72°C for 1 min. The final cycle was followed by an extension step at 72°C for 10 min. The PCR products were visualized on 2% agarose gels with 1x TRIS-acetate-EDTA buffer containing 200 ng/ml of ethidium bromide.

Genotyping was carried out on an ABI 3100 Avant automated DNA sequencer, with LIZ 500 (Applied Biosystems) as the internal lane standard (size standard). Sizing and allele calling was performed by using Genotyper version 2.0 software (Applied Biosystems). The allele data thus generated was used for further statistical analysis.

The data obtained on these 11 micro-satellite loci was subjected to analysis using POPGENE software (Yeh *et al*.

Table 1. Number of alleles (Na), effective number of alleles (Ne), F statistics, effective number of migrants (Nm) and heterozygosity statistics for each of 11 microsatellite loci

Locus	Na	Ne	F_{IS}	F_{IT}	F_{ST}	Nm	Obs_Het	Exp_Het	Nei
BMS2684	8	1.5327	0.5899	0.6452	0.1349	1.6038	0.1355	0.3478	0.3476
BMS2722	4	2.6863	0.0650	0.1000	0.0374	6.4271	0.5632	0.6282	0.6277
BMS2785	11	3.3022	0.2937	0.4628	0.2394	0.7942	0.3657	0.6977	0.6972
CSSM08	11	3.6034	0.0134	0.0720	0.0595	3.9543	0.6528	0.7231	0.7225
CSSM19	17	5.2992	0.0192	0.0968	0.0791	2.9100	0.7295	0.8119	0.8113
CSSM43	23	6.2592	0.3807	0.4751	0.1525	1.3894	0.4828	0.8410	0.8402
CSSM47	25	9.1294	0.0310	0.0852	0.0559	4.2222	0.8288	0.8912	0.8905
ILSTS05	14	3.0529	–0.1074	–0.0617	0.0412	5.8185	0.7200	0.6730	0.6724
ILSTS11	19	4.7483	0.2403	0.3244	0.1107	2.0082	0.544	0.7900	0.7894
ILSTS49	10	2.0085	0.1980	0.3609	0.2030	0.9812	0.3136	0.5025	0.5021
RM232	8	2.3742	0.3098	0.3691	0.0860	2.6566	0.3552	0.5793	0.5788
Mean	13.6364	3.9997	0.1548	0.244	0.1055	2.1195	0.5174	0.6805	0.68
SE	0.1893	0.0636					0.0059	0.00457	0.00456

1999) to obtain the expected and observed heterozygosity values for all loci. The F-statistics values F_{IS}, F_{IT}, F_{ST} were estimated and the number of migrants (Nm) was estimated using: $Nm = 0.25(1-F_{ST})/F_{ST.}$

Nei's genetic distance was estimated using POPGENE software. We performed Correspondence Analysis to understand the genetic relationship among the populations. The Correspondence Analysis and number of migrants per generation were calculated using the GENETIX software version 4.05 (Belkhir *et al*. 2004). The analysis of molecular variance (AMOVA) was carried out as implemented in ARLEQUIN version 3.1 software (Excoffier *et al*. 2005). The genetic distances among population were utilized to construct neighbor joining tree using software POPULATION 1.2.30 (Langella 1999) and tree was visualized using software TREEVIEW (Roderic Page 1996). The multivariate analysis for Principal Component was carried out using PCAGEN software (Goudet 2005). The deviation from the Hardy-Weinberg Equlibrium (HWE) for all locus-population combinations were determined using GENEPOP version 3.4 (Raymond and Rousset 1995).

The present study included 625 individuals of buffalo belonging to 34 districts of Uttar Pradesh. These were genotyped using 11 loci. All the heterologous loci were polymorphic in nature and were successfully amplified in buffaloes. The various genetic parameters estimated are given in Table 1. A total of 150 alleles were observed for the 11 microsatellite loci by genotyping 625 buffaloes. The number of alleles ranged from 4 to 25 for each loci. The locus BMS2722 with 4 alleles was least polymorphic while the locus CSSM47 with 25 alleles was highly polymorphic. The mean number of alleles was 13.64 (Table 1). However, most of the alleles were in low frequency, which is reflected by dramatic decrease in effective number of alleles. The mean effective number of alleles (4.00) represents the number of allele that cannot be lost due to chance. The estimates of F statistics F_{IS}, F_{IT} and F_{ST} represents inbreeding at the loci, total inbreeding and population differentiation respectively for each locus. The F_{IS} values were positive, with the exception of one locus (ILSTS05). The positive F_{IS} values mean local inbreeding in the buffalo population of Uttar Pradesh. The F_{ST} values were highest for BMS2785 (0.2394) and lowest for BMS 2722 (0.0374) but the mean F_{ST} for all loci was found to be 10.5%, based on allele frequency data. This signifies quite high differentiation between the buffaloes of different districts of Indo-Gangetic plains.

Table 1 shows the observed and expected heterozygosity values for each locus. The observed heterozygosity values for various loci ranged from 0.1355 (BMS2684) to 0.8288 (CSSM47). The high values for these loci make them suitable for diversity analysis. The mean observed heterozygosity value is 0.5174 with expected heterozygosity value of 0.6805. The values of observed heterozygosity are less than the expected values revealing population structure. The positive F_{IS} values also confirm the existence of local inbreeding effect which leads to population sub-structuring. The mean

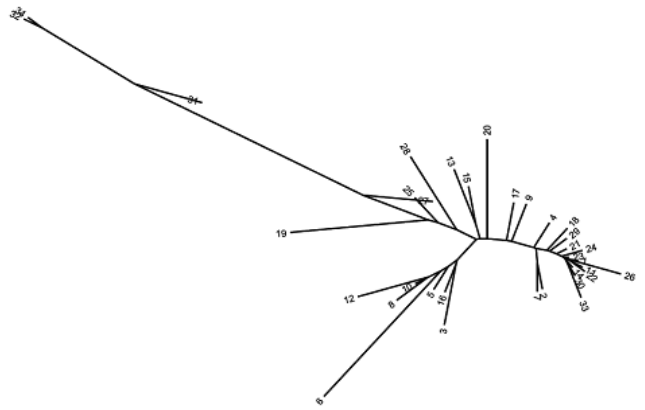

Fig. 1. Neighbour joining tree based on Nei genetic distance among different population distributed in Uttar Pradesh

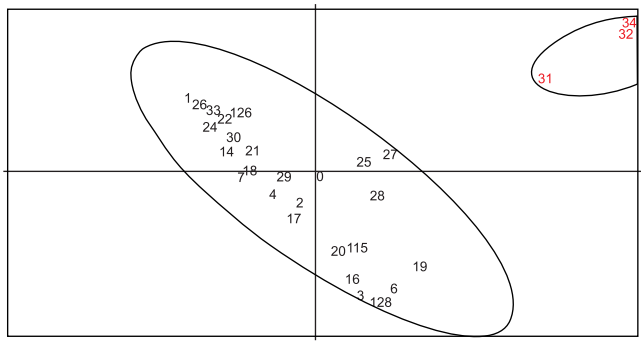

Fig. 2. Two dimensional Principal Component Analysis plot

Figs. 2–3. 1=Meerut, 2=Latlitpur, 3=EtahFirzMain, 4=AuriyaEtawah, 5=Rampur, 6=Badaun, 7=Jalaun, 8=Bareilly, 9=Farrukhabad, 10=Pilibhit, 11=Hardoi, 12=Hamirpur, 13=KnjKanpur, 14=Unnao, 15=Fatehpur, 16=LuckBaraban, 17=Raibareily, 18=Sitapur, 19=Allahabad, 20=Pratapgarh, 21=GondaBalram, 22=Faizabad, 23=AmbdkarNgr, 24=Mirazpur, 25=Jaunpur, 26=SiddhMaraj, 27=Azamgarh, 28Chandauli, 29=Santkbr Gorakh, 30=KrishiNgr, 31=Ghazipur, 32=Mau, 33=Deorria, 34=Balia.

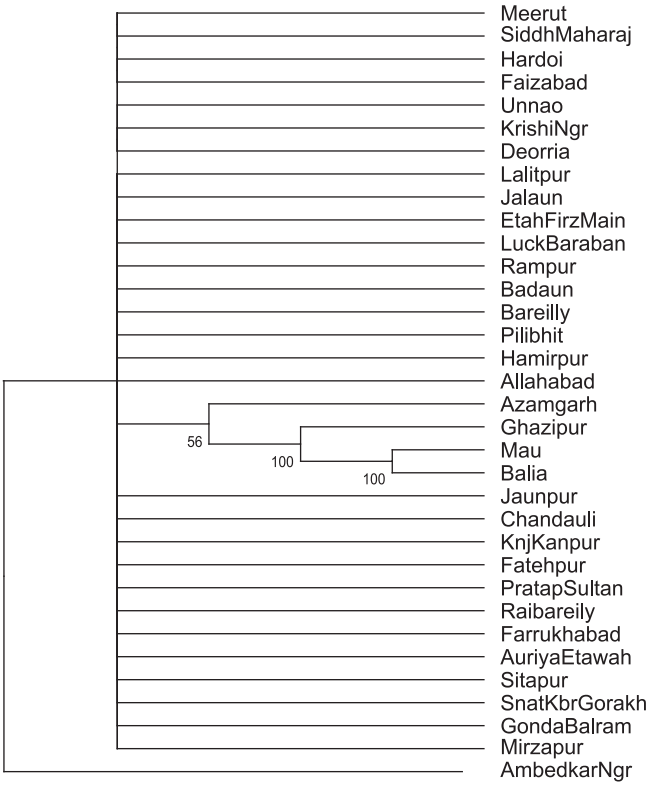

Fig. 3. Condensed topology tree of different population distributed in Uttar Pradesh

Nm (effective number of migrants) value among different districts is 2.11 which is quite high and exhibit geneflow among the districts. The Analysis of Molecular variance revealed that 7.2% variation among districts.

The deviation from Hardy Weinberg equilibrium (HWE)

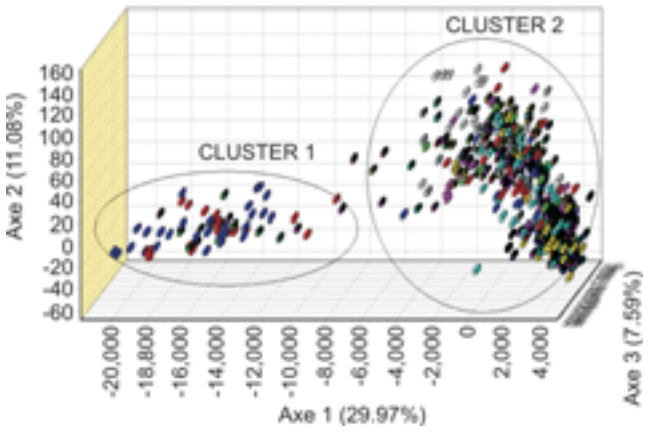

Fig. 4. Correspondence Analysis for clustering Population of Uttar Pradesh buffaloes

was estimated using GENEPOP version 3.4 (Raymond and Rousset 1995). The null hypothesis (Ho) of the population being in HWE was tested against alternative hypothesis (H_1) of heterozygotic deficit. The deviation was tested by calculating exact P values utilizing 10,000 dememorisation steps of 100 batches with 5000 iterations per batch. The four most frequent loci deviating from HWE were CSSM43 (18 districts), BMS2684 (13 districts), ILSTS11 (10 districts) and RM232 (8 districts). A total of 74 locus district combinations were not found to be in HWE.

The Nei's standard genetic distances were estimated and Neighbour joining algorithm was utilized to construct Neighbour joining tree. 100 bootstraps were performed over each locus. The Neighbour joining tree also showed one distinctive cluster while rest of the buffalo was in continuum forming a very large cluster Fig 1. The condensed topology of Neighbour joining tree (Fig 3) gave significant boot straps only for buffaloes of Mau, Balia and Ghazipur districts.

The Multivariate analysis using PCAGEN software (Goudet 2005) wherein we performed the Principal Component Analysis revealed the presence of two clear cut cluster as shown in the Fig 2. The buffalo populations of Balia, Mau and Ghazipur districts are quite distinctive from rest of the districts. There is in general a continuity of the buffalo population throughout the Indo-Gangetic plain which is reflected in the underlying genetic continuation as evidenced by microsatellite analysis. The other multivariate analysis also presents a similar trend as shown in Fig 4. Sub-structuring of the buffalo population is evidenced both in Correspondence Analysis and Principal Component Analysis but no clear demarcation is discernible except the two large clusters.

The mean numbers of alleles in the present study are quite high 13.64 which are higher as compared to the reports of (Vijh *et al.* 2008) for Bhadawari (7.18) and *Tarai* (8.00) which are also part of the present study though the animals and the loci differ. The numbers of alleles in a population are also property

of the locus itself and more so the present study has been on the basis of 11 micro-satellite loci. The observed and expected heterozygosity values in the present study are very similar to those reported by (Vijh *et al.* 2008) for various breeds of buffaloes of India. The F_{ST} value obtained in the present study is 7.2% which is less compared to 9.69% reported by (Vijh *et al.* 2008) and higher than reported by (Kumar *et al*. 2006) which was 3.4%. The values are thus between the two reported values and the reason can be that the buffaloes in Indo-Gangetic Plains are present in contiguity while in the earlier studies the samples were taken from only recognized breeds of the country with different geographical regions. The geographical distance among buffaloes is quite large from eastern Uttar Pradesh to western Uttar Pradesh and the differentiation of 7% may be due to upgradation of buffaloes by artificial insemination especially in the western part of the plains. In the present study, 3 districts of eastern Uttar Pradesh show clear cut differentiation to rest of the buffalo population of Indo-Gangetic Plains and thus may be genetically unique. This inference can be due to lack of infrastructure facilities and lack of Artificial insemination and the buffaloes may be true to the local population genetic structure without having much geneflow from nearby districts.

SUMMARY

The buffalo population of Uttar Pradesh constitutes 26.1% of the total buffalo population of India, yet this population has not been defined into distinct breeds or populations. The Indo- Gangetic plains have only one defined breed of buffalo named Bhadawari. Data on 11 micro-satellites was generated on 625 animals transversing the entire Indo-Gangetic basin. The mean number of alleles per locus was 13.34 while the effective number of alleles was 4.00 as most of the alleles were in a low frequency. The mean observed and expected heterozygosity were 0.52 and 0.68 respectively pointing towards the existence of population structure. The F_{IS} values were 0.15 which is statistically significant and points to local inbreeding in the buffalo population. The Analysis of molecular variance revealed 7.2% variation attributed to among district populations. The data revealed two distinctive clusters on the basis of correspondence analysis. The small cluster comprised of buffaloes of Mau, Balia and Ghazipur districts. The larger cluster also had discernible sub-structuring which could not be identified distinctively using multivariate analysis of Principal Component and Correspondence Anaysis.

ACKNOWLEDGEMENT

The financial support provided from NAIP ICAR, component 4 (C-1050) is gratefully acknowledged.

REFERENCES

Anonymous. 2003. *Livestock Census*. Department of Animal Husbandary and Dairying, Govt. of India, New Delhi, India.

Belkhir K, Borsa P, Chikhi L, Raufaste and Bonhomme F. 1996-2004. GENETIX 4.05, Windows TM software for population genetics. Laboratoire Genome, Populations, Interactions, CNRS UMR 5171, Universite Montpellier II, Montpellier (France). http: //www.univ-montp2.fr/~genetix/genetix/genetix.htm.

DAH. 2006. Department of Animal Husbandary and Dairying, Government of India, New Delhi, India.

Excoffier L, Laval G and Schneider S. 2005. An integrated software for population genetic data analysis. ARLEQUIN ver. 3.1. http: //lgb.unige.ch/arlequin/software.

Goudet J. 2005. PCAGEN: A computer program for Principal Component Analysis of gene frequency data. http: // www2.unil.ch/popgen/softwares/pcagen.htm.

Kumar S, Gupta J, Kumar N, Dikshit K, Navani N, Jain P and Nagarajan M. 2006. Genetic variation and relationships among eight Indian riverine buffalo breeds. *Molecular Ecology* **15**: 593–600.

Langella O. 1999. Population genetic software: Individuals or populations distances based on allelic frequencies, phylogenetic trees, file conversions. http: //www.bioinformatics.org/project/ ?group_id=84.

Navani N, Jain P K, Gupta S, Sisodia B S and Kumar S. 2002. A set of cattle microsatellite DNA markers for genome analysis of riverine buffalo (*Bubalus Bubalis*). *Animal Genetics* **33**: 149–54.

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

Page R D M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *A Computer Applications in the Biosciences* **12**: 357– 58. http: // taxonomy.zoology.gla.ac.uk/rod/treeview.html.

Raymond M and Rousset F. 1995. GENEPOP (version 3.4): Population genetics software for exact test and ecumenicism. *Journal of Heredity* **86**: 248–49. ftp: //ftp.cefe.cnrs-mop.fr/pub/ pc/msdos/genepop.

Sahai R and Vijh R K. 2000. *Domestic Animal Biodiversity-Conservation and Sustainable Management.* S.I. Publications, Karnal, India.

Vijh R K, Tantia M S, Mishra B and Bharani Kumar S T. 2008. Genetic Relationship and Diversity Analysis of Indian Buffalo. *Journal of Animal Sciencs* **86**: 1495–1502.

Vijh R K, Tantia M S, Mishra Bina, Bharani Kumar S T, Vij P K, Sirothia A R, Singh K P, Kharadi V B, Gajbiye P J, Iyue M and Gujjar B V. 2008. Genetic relationship among buffalo breeds of India. *Indian Journal of Animal Sciences* **78**(4): 378–83.

Yeh F, Boyle C, Rongcai T, Ye Z Y and Xian J M. 1999. POPGENE: ver 1.31. A Microsoft window based freeware for population genetic anaylsis, University of Alberta, Edmonton. http: // www.ualberta.ca/-fyeh/fyeh. Vice President, BAIF, Pune.