B-4974
[1-6]

RESEARCH ARTICLE Indian Journal of Animal Research

Elucidation of Genetic Divergence among Cattle Breeds of Tamil Nadu in Mitochondrial Genome

S. Vani, D. Balasubramanyam, S.M.K. Karthickeyan, M. Parthiban, P.S.L. Sesh

10.18805/IJAR.B-4974

ABSTRACT

Background: The present study was carried out for genetic characterization and assessment of evolutionary status of native cattle breeds of Tamil Nadu, India.

Methods: Complete mitochondrial genome sequence of 15 pooled samples belonging to five cattle breeds of Tamil Nadu was carried out for the first time using Illumina platform. These mitogenomes were utilized in the present investigation to study the mitochondrial diversity. using suitable statistical tools.

Result: The sequences encompassed 16,338 to 16,340 nucleotides with 273 variants. Low degree of genetic divergence and polymorphism was observed in the population. Higher rate of migrants (11.77) between populations caused increase in gene flow. F_{ST} value of -0.04435 revealed low level of genetic differentiation between cattle breeds of Tamil Nadu. A significant negative Tajima's D ($P<0.05$) for all the *Bos indicus* cattle indicates operation of purifying selection and population expansion among the genetic groups under study. Maximum likelihood reconstruction using complete mitochondrial sequences of the present study combined with previously reported sequences of Nellore cattle and eight *Bos taurus* cattle breeds illustrated close genetic relationship among the *Bos indicus* cattle under study with a clear demarcation from the *Bos taurus* cattle indicating higher genetic divergence between the two lineages. This finding was also supported by multi-dimensional scaling. Analysis of molecular variance revealed no differentiation between cattle breeds of Tamil Nadu.

Key words: *Bos indicus* cattle, Genetic divergence, Mitogenome, Phylogeny, Whole genome sequencing.

INTRODUCTION

Archaeological and genetic evidences suggest that modern cattle might have arose from several domestication events of aurochs (*Bos primigenius*) and resulted in the formation of *Bos indicus* and *Bos taurus* lineages around 8,000-10,000 years ago (Loftus *et al*. 1994). India has large indigenous cattle population with 50 distinct breeds. Extensive and different range of agro-ecological zones in India has assisted in the development of each breed (Sharma *et al.* 2015) with distinct characteristics. Indian cattle breeds are broadly categorized into dairy, draft and dual purpose breeds based on their utility. However, majority of them are draft breeds, under severe neglect resulting in continuous decline of their germplasm. The introduction of highly productive exotic breeds and demographic pressure are also contributing to the loss of precious traits and reduction in population size of indigenous breeds. Cattle breeds of Tamil Nadu (Alambadi, Bargur, Kangayam, Pulikulam and Umblachery) are characterized as tropical/subtropical breeds possessing heat tolerance and parasite resistance and can survive in a harsh environment and on low-quality roughages.

Thus, assessing the genetic characteristics of such indigenous breeds would be helpful in better understanding of genetic diversity as well as formulating action plans for conservation and management. The genetic diversity of cattle breeds of Tamil Nadu has previously been investigated using a range of techniques *viz*. karyotypic analysis (Parameswari *et al*. 2019), microsatellite DNA markers (Barani *et al*. 2015) and Y-chromosome specific microsatellite polymorphisms (Jeevan, 2022). Maternally inherited mitochondrial DNA (mtDNA) has been extensively used to determine the genetic variation and phylogenetic relationships of domestic cattle, but related studies have been mostly restricted to the short hypervariable region which makes it impossible to clearly distinguish between some important ancient branches within the phylogenetic tree. To date, most mtDNA studies have focused on the control region (D-loop) (Jakaria *et al*. 2019). Therefore, the present study aims in including the additional information on the matrilineal genetic diversity and phylogenetic status of the cattle breeds of Tamil Nadu using mitochondrial DNA (mtDNA) sequence data obtained from the whole genome sequencing. Using the entire mitochondrial genome sequence to study the genetic structure of animals would provide us refined phylogenies of maternal lineages.

Department of Animal Genetics and Breeding, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai-600 007, Tamil Nadu, India.

Corresponding Author: S. Vani, Department of Animal Genetics and Breeding, Madras Veterinary College, Tamil Nadu Veterinary and Animal Sciences University, Chennai-600 007, Tamil Nadu, India. Email: vanireddy12786@gmail.com

How to cite this article: Vani, S., Balasubramanyam, D., Karthickeyan, S.M.K., Parthiban, M. and Sesh, P.S.L. (2022). Elucidation of Genetic Divergence among Cattle Breeds of Tamil Nadu in Mitochondrial Genome. Indian Journal of Animal Research. DOI: 10.18805/IJAR.B-4974.

Submitted: 01-07-2022 **Accepted:** 31-08-2022 **Online:** 15-09-2022

MATERIALS AND METHODS

Sample collection and processing

The present study was carried out at department of Animal Genetics and Breeding, Madras Veterinary College, Chennai during 2021-2022. A total of 302 blood samples were collected from four registered cattle breeds *viz*. Bargur, Kangayam, Pulikulam, umblachery and one unregistered cattle (Alambadi) population maintained at their respective cattle research stations in different agro-climatic regions of Tamil Nadu, India. Genomic DNA was extracted using phenol-chloroform method (Sambrook *et al*. 1989) with slight modifications by adding DNAzol solution instead of proteinase-k. After assessing the quality and quantity of genomic DNA, a total of 79 animals from five breeds were selected based on the sex of the animal and available milk records. They were categorized and pooled into three groups from each breed as bulls, medium yielding dams (>2 kg milk yield per day) and low yielding dams (<1 kg milk yield per day). Thus a total of 15 pooled samples (Alambadi: ACG1, ACG2 and ACG3; Bargur: BCG1, BCG2 and BCG3; Kangayam: KCG1, KCG2 and KCG3; Pulikulam: PCG1, PCG2 and PCG3; Umblachery: UCG1, UCG2 and UCG3) were prepared sent for whole genome sequencing.

Illumina sequencing and reconstruction of mitochondrial genomes

The genomic DNA was sheared into fragments and genome sequence library was constructed by ligating specialized adapters at both ends of sheared fragments using NEBNext® Ultra™ II FS DNA Library Prep kit. The libraries were subjected to sequencing on Illumina HiSeq 2500 and Novoseq 6000 platforms (Illumina, San Diego, CA) and paired end reads were generated. After sequencing, the filtered reads were mapped against *Bos indicus* complete mitochondrion reference sequence, (Nellore breed, Genbank accession no. GCF-000247795.1) using BWA-MEM algorithm v0.7.17-r1188 with the default parameters. Variant calling of mitochondrial genes was performed using Genome Analysis Tool Kit Haplotype Caller v 4.2.0.0-1.

Data analysis

Polymorphism and divergence analyses of the sequences were performed using DnaSP v 6.12.03 (DNA Sequence Polymorphism) software (Rozas *et al*. 2017). The polymorphism was measured in terms of nucleotide diversity, Pi (π) which gives the average number of nucleotide differences between two sequences (Nei, 1987).

The divergence was measured in terms for average number of nucleotide differences (Tajima, 1983) at individual sample level and the total data, number of haplotypes 'H' (excluding gaps in the alignment), haplotype diversity 'HD' (Nei, 1987).The genetic differentiation was measured by various estimates like haplotype based H_S, H_{ST}, nucleotide sequenced based K_S, K_{ST} and Z (Hudson *et al*. 1992) and Snn (Hudson, 2000). The statistical methods for testing the hypothesis of genetic differentiation among the breeds were given by Chi-square test for haplotype data. (Nei, 1987; Hu dson *et al*. 1992) and permutation test for the various measures (Hudson *et al*. 1992) were utilized. The gene flow estimates were computed assuming haploid data as the mitochondrial DNA is haploid. The estimates were obtained from nucleotide sequence data as F_{ST} (Lynch and Crease, 1990) with the estimated number of migrants (N_m).

The conservation of sequences among the populations was analyzed using sliding window analysis with default parameters in DnaSP. The conservation was measured in terms of conservation index (C) and Homozygosity (H); 'C' as a proportion of conserved sequences and 'H' as 1-heterozygosity.

To test the theory of neutral evolution, the test statistics such as Tajimas's D (Tajima, 1989), Fu and Li's D and Fu and Li's F (Fu and Li, 1993; Fu, 1997) were computed using proportion of segregating sites within a gene by DnaSP software.

Mitochondrial DNA sequences of studied breeds were compared with published mitogenomes belonging to Nellore cattle and *Bos taurus* cattle breeds from NCBI database. All the mitochondrial sequences were aligned using Muscle algorithm available in MEGA x (Kumar *et al*. 2018). Phylogenetic analysis was performed using MEGA x (Molecular Evolutionary Genetic Analysis) version 11. Model Test (Posada and Crandall 1998) was performed to check the best suitable model that explains the sequence divergence based on AICC and BIC values. Neighbour joining tree was constructed using *Bos taurus* mitochondrial sequences as an out group in MEGA X software. The topology was tested using bootstrap approach with 1000 replicates. Multi-dimensional scaling was plotted to know the clustering of *Bos indicus* and *Bos taurus* cattle breeds using SPSS v 20 software. Analysis of molecular variance among the populations was done using AMOVA in Genalex v 6.5.

RESULTS AND DISCUSSION

Whole genome sequencing

For each sample, the depth of coverage for mitochondrial genome ranged from 145.82 to 1652.72 x with an average of 802.59 x when the paired-end reads generated were aligned to the *Bos indicus* reference genome sequence (GCF-000247795.1-*Bos-indicus*-1.0). The mitochondrial genome sequences ranged from 16,338 to 16,340 bp with 13 protein-coding genes, two ribosomal RNA (12S and 16S rRNA) genes, 22 transfer RNA (tRNA) genes and one control region of 913 bp (D-loop) in cattle breeds of Tamil Nadu as found in other mammals.

Variant calling

A total of 273 variants observed across 13 protein coding regions from the 15 pooled samples sequenced; of which, 268 were bi-allelic SNPs, three multi-allelic SNPs and two InDels. Pooling of the samples strenghtened the richness of variant calling in the present study as also supported by Gautier *et al*. (2013). Of the biallelic SNPs, 256 SNPs were observed to be transitions while the remaining 12 were transversions. Out of total variants detected, 157 variants

were found to be synonymus mutations with very low effect; whereas, 30 variants were observed to be non synonymous (missense) mutations Eighty six variants (84 SNPs and 2 InDels) were detected on the upstream region and have a modifier effect on the protein synthesis. List of variants observed in all the 13 protein coding regions were presented in Table 1. Highest number of variants was found in Pulikulam cattle across the coding regions. ND1 gene was observed to have highest number of variants (71) across 15 samples but is restricted to 45 loci indicating its polymorphism as 1.57 per locus. Lowest number of variants (5) were observed in ATP8 (five loci) and ND4L (four loci) genes. Information pertaining to number of polymorphic loci pertaining to 13 protein coding genes in five cattle breeds of Tamil Nadu is shown in Fig 1.

Genetic structure of indigenous cattle breeds

The amount of genetic variation within a population provides an insight into the demographic structure and evolutionary history of a population. The polymorphism and genetic diversity indices of five cattle breeds were detected using DnaSP v 6.12.03 software for the 15 mitochondrial DNA sequences and are presented in Table 2. The polymorphism, measured in terms of nucleotide diversity (π) is the average number of nucleotide differences between two sequences. Moderate amount of polymorphism was observed in Alambadi, Bargur, Kangayam, Pulikulam and Umblachery cattle breeds with average difference in nucleotides between the two sequences within the populations as 0.11, 0.12, 0.14, 0.76 and 0.12 per cent respectively. A total of 390 sites were found to be polymorphic of which, 46, 58, 43, 205 and 38

Table 1: Total number of variants observed in the protein coding region of Mt genome among five cattle breeds.

Gene	No. of SNPs	Substitutions			Effect/Location of SNP			No. of In Dels	Total
		Transition	Transversion	Mixed	Synonymus	Non synonymus	Upstream		
ATP6	8	7	1	-	6	2	-	-	8
ATP8	5	5	-	-	2	2	1	-	5
COX1	26	24	1	1	21	-	5	-	26
COX2	8	8	-	-	5	-	3	-	8
COX3	16	16	-	-	11	5	-	-	16
CYTB	19	19	-	-	15	4	-	-	19
ND1	69	63	5	1	16	3	50	2	71
ND2	15	15	-	-	13	1	1	-	15
ND3	10	9	1	-	8	1	1	-	10
ND4	20	19	1	-	18	2	-	-	20
ND4L	5	5	-	-	4	-	1	-	5
ND5	39	35	3	1	29	10	-	-	39
ND6	31	31	-	-	9	-	22	-	31
Total	271	256	12	3	157	30	84	2	273

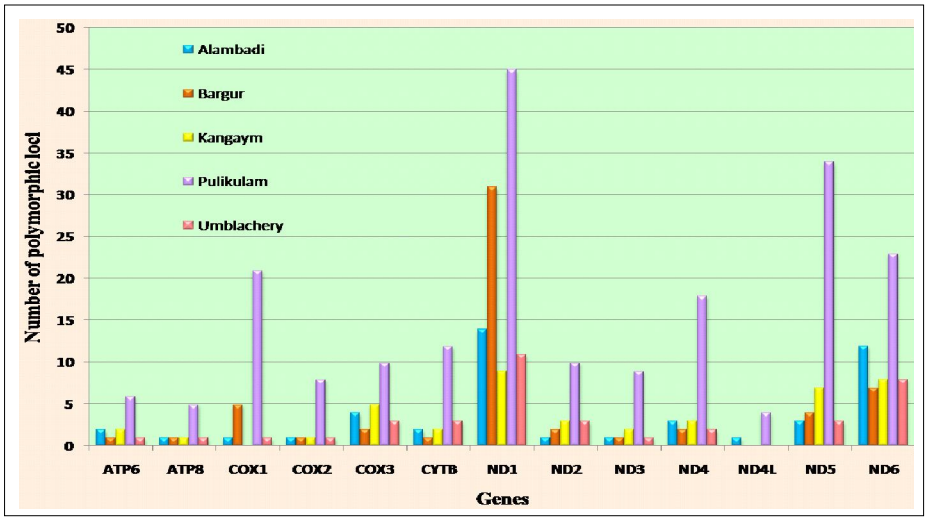

Fig 1: Breed-wise number of polymorphic loci in mitochondrial protein coding genes of cattle breeds of Tamil Nadu.

were observed in Alambadi, Bargur, Kangayam, Pulikulam and Umblachery cattle breeds respectively. Of these 390 sites, 304 sites were found as segregating. Number of polymorphic sites in the present study are more (390) than that revealed by Chung (2013) in Korean cattle (286-288). A total of 22 regions in Alambadi, Kangayam and Umblachery; 23 regions in Pulikulam and 25 regions in Bargur cattle are found to be conserved significantly ($p<0.05$).

Genetic diversity has a vital role in the survivability and adaptation of the populations. The average number of nucleotide differences between the two sequences (ð) within the populations and the number of segregating sites revealed high haplotype diversity. All the15 sequences of five cattle breeds were categorised into 15 different haplotypes with haplotype diversity (HD) of 1.0 indicating high degree of haplotype diversity. Haplotype diversity of the present study was more than that reported by Chung (2013) in Korean cattle (0.052 to 0.20) and Petretto *et al*. (2022) in Sardinian local cattle stock (0.879). The genetic differences were observed between populations of the same breed, indicating that the identified haplotypes may be used to characterize group specificities of each cattle breed (bulls, moderate yielders, low yielders).

Genetic differentiation

DnaSP was used to compute nucleotide test statistics such as Ks, Kst , Snn and haplotype statistics such as Hs and Hst to describe the genetic divergence within the populations. All the 15 samples were separated into 15 haplogroups. The Hs, Hst, Ks, Kst, Z and Snn values observed in the present study were 1.00, 0.00, 45.00, -0.032, 57.20 and 0.033 respectively. As per overall genetic differentiation metrics, it was observed that no genetic differentiation among the *Bos indicus* cattle breeds of Tamil Nadu under study. High gene flow between the types, which could be attributable to introgression during breeding programmes and subsequent selection, could explain the reason for no genetic distinction.

Table 2: Polymorphism and diversity indices of mitochondrial genome sequences among different cattle breeds.

Breed	Number of polymorphic sites	Nucleotide diversity, Pi (π)	No. of segregating sites (S)	Average number of nucleotide differences (K)	Tajima's D
Alambadi	46	0.00113	27	18	-2.095*
Bargur	58	0.00125	30	20	
Kangayam	43	0.00142	34	22.66	
Pulikulam	205	0.0076	182	121.33	
Umblachery	38	0.00129	31	20.66	
Total	390	0.00261	304	42.63	

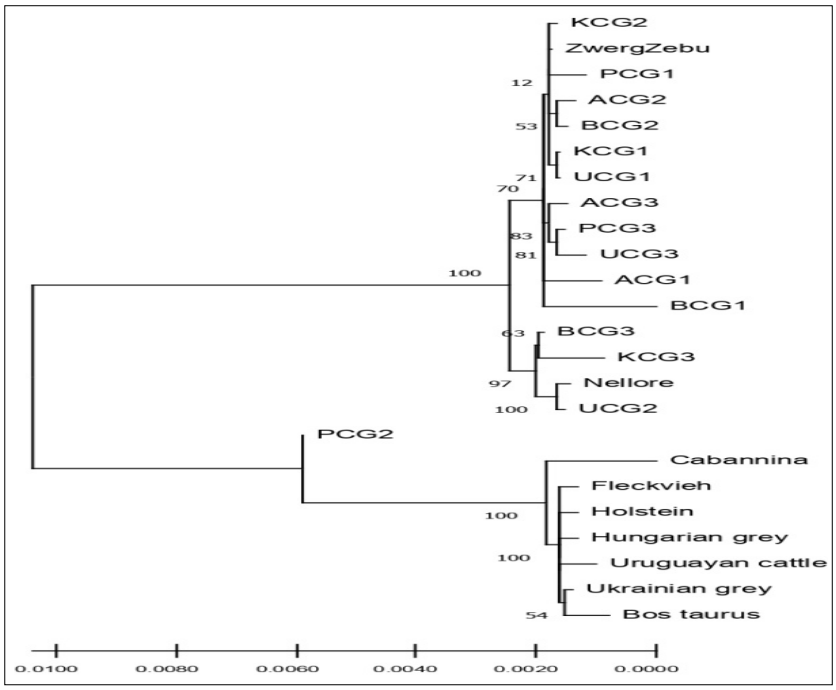

Fig 2: Phylogenetic analysis of cattle, based on the complete mitochondrial DNA sequence. Accession numbers of other breeds used are as follows: Zwerg Zebu-AF492350; Nellore-NC005971; Holstein-DQ124418; Ukrainian Grey-GQ129208; Hungarian Grey-GQ129207; Fleckvieh-AF492351: Cabannina- EU177867.

Gene flow

Gene flow is a crucial technique for spreading genetic variation between populations. The overall F_{ST} value was 0.4932 indicating genetic differentiation among the populations analysed when 24 sequences from *Bos indicus* and *Bos taurus* cattle were included together. This indicates that these populations are genetically distinct. On the contrary, the overall F_{ST} was -0.04435 with a net migration rate of 11.77 when only *Bos indicus* sequences of Tamil Nadu were included (three each from Alambadi, Bargur, Kangayam, Pulikulam and Umblachery) indicating high rate of intermixing among the five populations and the effective migrants were also similarly high (11.77), suggesting high gene flow between them. Gene flow at high rates can minimise genetic differentiation between the two populations, resulting in increased homogeneity. Because the majority of cattle breeding in India have not been systematically evolved, gene movement between populations is generally expected. F_{ST} value observed in the present study (value) was lower than that reported by Petretto *et al*. (2022) in Sardinian local cattle (0.056) with high amount of genetic differentiation.

Neutral evolution

The test of neutral evolution analyzed based on the total number of mutations and segregating sites across all 15 sequences of *Bos indicus* breeds, revealed statistically significant (P<0.05) negative values for Tajimas's D (-2.095), Fu and Li's D (-2.974,) and Fu and Li's F (-3.150) test statistics. A significant negative Tajima's D (P<0.05) for all the *Bos indicus* cattle indicated an excess of low frequency polymorphisms than expected, indicating existence of purifying selection and population expansion among the genetic groups under study. Similar findings were reported by Petretto *et al*. (2022) in Sardinian local cattle stock.

Phylogenetic analysis

Maximum likelihood based phylogenetic analysis was performed with 15 complete mitogenomes of five breeds combined with eight taurine sequences and one Nellore cattle reference mitogenome sequence and is shown in Fig 2. HKY+G+I (Hasegawa-Kishino-Yano + Gamma distribution + Evolutionarily invariable sites) model was found to have lowest AICC and BIC values and was selected as the best model to explain the divergence of sequences. The phylogenetic tree divided all the 24 sequences into two distinct haplogroups: *Bos indicus* and *Bos taurus* (Fig 1). It divided *Bos indicus* clade into two subclades. Interestingly, Zwerg Zebu, an European dwarf zebu cattle breed was placed in the *Bos indicus* group. Similar findings were observed by Pramod *et al.* (2018) where, Zwerg Zebu cattle was placed in the *Bos indicus* lineage. It was also found that one sequence from Pulikulam breed of cattle (PCG2) was unambiguously associated with the *Bos taurus* haplogroup, but representing an unknown divergent mitochondrial sub-haplogroup.

Multi-dimensional scaling

Multi dimensional scaling clearly segregated 24 sequences into two different clusters as *Bos indicus* and *Bos taurus* cattle. These results were in accordance with the phylogenetic analysis. The Zwerg Zebu cattle was placed in *Bos indicus* cluster whereas one sequence from Pulikulam cattle to be joined to *Bos taurus* cluster. The MDS plot of 24 mitochondrial sequences belonging to *Bos indicus* and *Bos taurus* cattle is displayed in Fig 3.

Analysis of molecular variance (AMOVA)

AMOVA analysis indicated there was no differentiation of the five cattle breeds of Tamil Nadu (between-population component of variation is zero) although all the diversity was gathered at the within-breed level (Table 3).

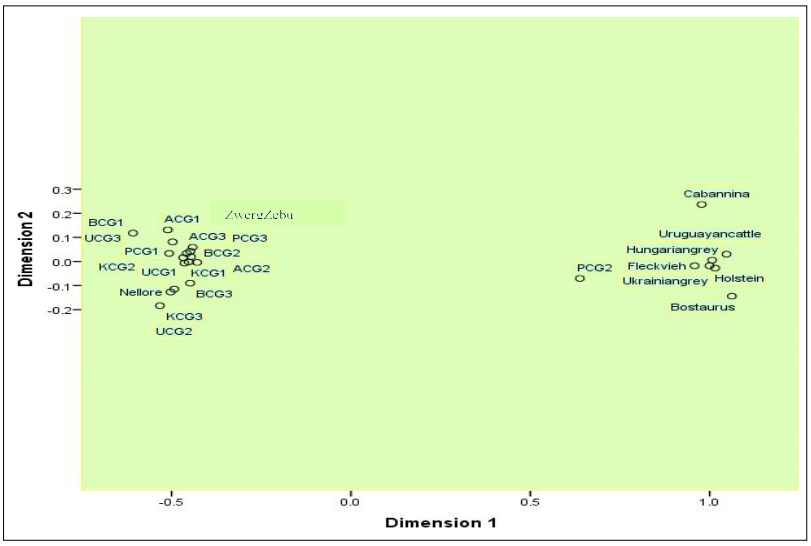

Fig 3: MDS plot showing clustering of *Bos indicus* and *Bos taurus* cattle breeds.

Table 3: Analysis of molecular variance (AMOVA) of mtDNA sequences from cattle breeds of Tamil Nadu.

Source of variation	df	Sum of squares	Mean sum of squares	Percentage of variation
Among populations	4	93.533	23.383	0
Within population	10	250.333	25.033	100

CONCLUSION

Mitochondrial DNA (mtDNA) analysis is a decisive tool in assessing the maternal origin, phylogeny and population structure of domestic animals. Cattle breeds of Tamil Nadu are known for heat tolerance and disease resistance, and can survive in a harsh environment and on low-quality roughages. In conclusion, our findings revealed no genetic differentiation among the *Bos indicus* cattle breeds of Tamil Nadu which might be due to high gene flow between them. Hence, assessing the genetic structure of these cattle was of meticulous importance for designing breeding strategies and conservation programs in future.

ACKNOWLEDGEMENT

The authors are thankful to GoI-NPBB-TNLDA Project on "Establishment of National Facility for Molecular Screening of Cattle for Inherited Disorders and Parentage Verification", functioning at the Dept. of Animal Genetics and Breeding, Madras veterinary college, Chennai. India.

Conflict of interest: None.

REFERENCES

Barani, A., Rahumathulla, P.S., Rajendran, R., Kumarasamy, P., Ganapathi, P., Radha, P. (2015). Molecular characterization of Pulikulam cattle using microsatellite markers. Indian Journal of Animal Research. 49(1): 36-39.

Chung, H. (2013). Phylogenetic analysis and characterization of mitochondrial DNA for Korean native cattle. Open Journal of Genetics. 3(1): 12-23.

Fu, Y. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 147: 915-925.

Fu, Y.X. and Li, W.H. (1993). Statistical tests of neutrality of mutations. Genetics. 133: 693-709.

Gautier, M., Foucaud, J. Gharbi, K. Cezard, T. Galan, M. Loiseau, A. Thomson, M. Pudlo, P. Kerdelhue, C., Estoup, A. (2013). Estimation of population allele frequencies from next-generation sequencing data: Pool versus individual based genotyping. Molecular Ecology. 22(14): 3766-3779.

Hudson, R.R., Boos, D.D., Kalpan, N.L. (1992). A statistical test for detecting population subdivision. Molecular Biology and Evolution. 9: 138-151.

Hudson, R.R. (2000). A new statistic for detecting genetic differentiation. Genetics.155: 2011-2014.

Jakaria, J., Musyaddad, T., Rahayu, S., Muladno, M., Sumantri, C. (2019). Diversity of D-loop mitochondrial DNA (mtDNA) sequence in Bali and Sumba Ongole cattle breeds. Journal of the Indonesian Tropical Animal Agriculture. 44(4): 335-345.

Jeevan, C. (2022). Prediction of semen production potential using conventional and machine learning approaches. Thesis submitted to Tamil Nadu Veterinary and Animal Sciences University.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution. 35: 1547-1549.

Loftus, R.T., MacHugh, D.E. Bradley, D.G., Sharp, P.M., Cunningham, P. (1994). Evidence for two independent domestications of cattle. Proceedings of the National Academy of Sciences. 91(7): 2757-2761.

Lynch, M. and Crease, T.J. (1990). The analysis of population survey data on DNA sequence variation. Molecular Biology and Evolution. 7: 377-394.

Nei, M. (1987). Molecular Evolutionary Genetics. Columbia University. Press, New York.

Parameswari, S., Cauveri, D., Karthickeyan, S.M.K., Arunachalam. K., Kumarasamy. P. (2019). Cytogenetic Characterisation of Alambadi breed of cattle in Tamil Nadu. Indian Veterinary Journal. 96(5): 49-52.

Petretto, E., Dettori, M.L., Pazzola, M., Manca, F., Amills, M., Vacca, G.M. (2022). Mitochondrial DNA diversity of the Sardinian local cattle stock. Scientific Reports. 12(1): 1-7.

Pramod, R.K., Velayutham, D., Sajesh, P.K., Beena, P.S., Zachariah, A., Zachariah. A., Chandramohan, B., *et al.* (2018). The complete mitochondrial genome of Indian cattle (*Bos indicus*). Mitochondrial DNA B: Resources. 3(1): 207-208.

Posada, D. and Crandall, K.A. (1998). Model test: Testing the model of DNA substitution. Bioinformatics. 14: 817-818.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution. 34(12): 3299-3302.

Sambrook, J. (1989). Isolation of DNA from mammalian cells: Protocol I. Molecular cloning. A Laboratory Manual. Pp. 916-919.

Sharma, R., Kishore, A., Mukesh, M., Ahlawat, S., Maitra, A., Pandey, A.K., Tantia, M.S. (2015). Genetic diversity and relationship of Indian cattle inferred from microsatellite and mitochondrial DNA markers. BMC Genetics. 16(1): 1-12.

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

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123: 585-595.