PRINT: ISSN 0972-3757 ONLINE: 2456-6360
Int J Hum Genet, 15(2): 81-87 (2015)
DOI: 10.31901/24566330.2015/15.02.05

Alu Insertion/Deletion Polymorphism in Four Tribes of South India

A. Krishnaveni* and K. Prabhakaran

PG & Research Department of Zoology, Periyar E. V. R. College (Autonomous), Tiruchirappalli 620 023, Tamil Nadu, India
**E-mail: krishnaveni.balakris@gmail.com*

KEYWORDS Dravidian. Genome Diversity. Heterozygosity. Genetic Distance. Autosomal Markers

ABSTRACT The present study was conducted to investigate the genetic diversity and affinities of four South Indian tribal populations namely, Malaikuravan, Malasar, Palliyan and Kattunaikkan. DNA samples from 184 unrelated individuals belonging to four tribal populations were analyzed for nine human-specific insertion and deletion polymorphic loci. The results indicate that all the studied biallelic loci are highly polymorphic in terms of allele frequencies and average heterozygosities (H_T = 0.49) in all the study populations. The genomic diversity (G_{ST}) of the four tribal populations was quite low (2.4%). The neighbor joining tree and centroid analysis showed that the four tribal populations under consideration are equally distant from other Indian tribal and world populations. The results reveal that the study populations have received moderate gene flow compared to other populations.

INTRODUCTION

The ethnic population of India is highly diverse both, biologically and culturally (Majumder 1998). The structure of the Indian population consists of tribal and non-tribal populations. The tribes are considered to be the original inhabitants of India (Ray 1973), and form 8.2 percent of the total population (Census of India 2001). The genetic studies on Indian populations has pointed towards the firm social boundaries, strict endogamy practices and evolutionary forces that have played a major role in building the diverse genetic structure of the present day Indian populations (Singh et al. 2012). In South India, a different pattern of allele and haplotype frequencies and mean values are generally observed among the tribal population groups, for which one of the main causes might be the small population sizes. Inbreeding is prevalent among certain communities and in most of the different population groups particularly from South India, which might have also resulted in the marked variation in distribution of frequencies, mean values of different genetic markers and morphological traits (Bhasin 2009).

Address for correspondence:
A. Krishnaveni,
PhD Research Scholar,
PG and Research Department of Zoology,
Periyar E. V. R. College (Autonomous),
Tiruchirappalli 620 023, Tamil Nadu, India
Mobile: +91-9487438628
Fax: +91 431-2332010
E-mail: krishnaveni.balakris@gmail.com

Transposable Elements (TEs) are powerful drivers of evolution. TEs constitute the majority of genomic DNA in many eukaryotes, and they dramatically shape genetic content by causing mutations, rearrangements, and sequence duplications. Of increasing significance is the link between these transposon-mediated mutations and disease (Singh et al. 2014). Alu (member of a SINE family) insertion/deletion polymorphisms offer several advantages over other nuclear DNA polymorphisms for human evolutionary studies. They are rapid, simple and stable with newly inserted elements and rarely undergo deletion. It is recently proved that the ACE deletion/deletion polymorphism could affect the athletic ability in Turkish population (Inanir et al. 2014). The present study aimed to investigate the genomic diversity, genetic differentiation and genomic affinities of the four South Indian tribal populations, based on nine human-specific Alu insertion/deletion polymorphisms in the nuclear genome.

MATERIAL AND METHODS

Blood samples (5-10 ml veniepuncture) were drawn from 184 healthy, unrelated adult volunteers from four endogamous Dravidian speaking tribal populations of South India with prior informed consent. The tribal groups are confined to the villages of hilly tracts and valleys of four different districts of Tamil Nadu, India. They include Malaikuravan from Thiruvannamalai

(n=47), Malasar from Anaimalai in Coimbatore (n=46), Palliyan from Theni (n=50) and Kattunaikkan from Chennai (n=41). The genomic DNA was extracted from whole blood using the salting out procedure and the non-enzymatic method, and was suspended in 10mM Tris and 0.1mM EDTA for genotyping. The nine polymorphic loci were genotyped by a standard 30-cycle PCR. Appropriate annealing temperatures and additives were optimized for each system. The PCR protocols followed for the present study have been reported previously (Stoneking et al. 1997; Majumder et al. 1999). After PCR amplification, amplicons were separated by electrophoresis. Later the EtBr stained gel was visualized under UV and documented.

Statistical Analysis

Allele frequencies at each of the nine loci were calculated by direct gene counting method and genotypes were analyzed for Hardy-Weinberg Equilibrium (HWE). Nei's expected heterozygosity (Nei 1973) of population was analyzed using the Pop Gene version 1.32 software (Yeh et al. 1999). To assess the extent of gene differentiation among the population groups, Nei's measure of gene diversity was calculated separately for each locus and all loci considered jointly. The genomic affinity among the populations was assessed by dendrograms, which were constructed by the neighbor joining (NJ) tree (for 1000 replicates) method using the DISPAN program (Ota 1993). The tree reconstruction methods depict population relationships as a series of bifurcations, which are commonly interpreted as population split; however, it is important to realize that clusters of populations in such trees could arise from migration instead of shared ancestry. The Harpending and Ward (1982) model was followed to understand the pattern of gene flow.

RESULTS

Allele Frequencies and Genomic Diversity within Populations

The number of chromosomes examined and allele frequencies of Alu FXIIIB, ACE, PV92, TPA25, APO, PLAT, D1, mt NUC and CD4 deletion locus are presented in Table 1 separately for the four tribal populations of South India. The (-) allele frequency for CD4 del locus is presented because the deletion allele is human specific. In most of the cases the heterozygosity for a biallelic marker has attained the maximum attainable value (0.5). The Hardy-Weinberg equilibrium was tested using the χ^2 test for goodness of fit after bonferroni correction. It was observed that eighteen values are significant at a five percent level. The heterozygosities at each locus and the average heterozygosities for overall loci for the study population are given in Table 2. The average heterozygosity ranges from 0.475 (Malasar) to 0.491 (Palliyan). The values of average heterozygosities are almost closer to the theoretical expectation of HW equilibrium.

Genetic Diversity Between Populations

To quantify the amount of genetic differentiation among inter-population, the G_{ST} values (a measure of the inter-population variability) for

Table 1: Allele frequencies and Hardy-Weinberg ÷² goodness of fit values at nine indel polymorphic loci in four south Indian tribal populations

Locus	*Population*											
	Malaikuravan			*Malasar*			*Palliyan*			*Kattunayakkan*		
	n[a]	*P(+)*	÷²	*n*[a]	*P(+)*	÷²	*n*[a]	*P(+)*	÷²	*n*[a]	*P(+)*	÷²
Alu FXIIIB	80	0.5750	0.270	68	0.5588	5.786*	84	0.5476	11.634*	78	0.3846	12.01*
Alu ACE	94	0.6915	4..30	72	0.5694	12.12	86	0.4884	8.890*	78	0.4615	17.784*
Alu TPA25	88	0.5795	0.83	92	0.4891	4.113	100	0.3800	12.02*	82	0.5976	8.016*
Alu mtNUC	90	0.4778	0.420	90	0.5222	0.598	84	0.5357	24.920*	82	0.4268	17.380*
Alu PV92	94	0.6489	5.712*	80	0.5875	4.260	78	0.5513	1.936	80	0.6000	8.984*
Alu APO	92	0.6196	5.514*	72	0.5833	6.416*	88	0.3864	2.786	82	0.5000	5.488*
Alu PLAT	94	0.5426	0.889	84	0.5833	18.60*	82	0.5610	10.450*	82	0.4756	0.2065
Alu D1	90	0.5000	0.610	92	0.7826	3.420	68	0.5147	1.580	60	0.4167	2.970
Alu CD4(del)	88	0.4773	1.752	88	0.5000	6.470*	100	0.4800	4.384	72	0.4167	14.260*

* significant at 5% level

Table 2: Heterozygosities at individual locus and average hetrozygosity based on nine polymorphic loci in each four south Indian tribal populations

Locus	*Malai-kuravan*	*Malasar*	*Palliyan*	*Kattuna-ikkan*
Alu FXIIIB	0.4888	0.4931	0.4955	0.4734
Alu ACE	0.4867	0.4904	0.4997	0.4970
Alu TPA25	0.4873	0.4998	0.4712	0.4810
Alu mtNUC	0.4990	0.4990	0.4974	0.4893
Alu PV92	0.4556	0.4847	0.4947	0.4800
Alu APO	0.4714	0.4861	0.4742	0.5000
Alu PLAT	0.4964	0.4861	0.4926	0.4988
Alu D1	0.5000	0.3403	0.4996	0.4861
Alu CD4(del)	0.4990	0.5000	0.4992	0.4861
All loci	0.4805	0.4755	0.4916	0.4839

each polymorphic locus were determined. The result of the gene diversity analysis is presented in Table 3. The G_{ST} values calculated for each locus, to determine the degree of gene differentiation between the populations, varies between 0.39 percent at the CD4 locus to 7.6 percent at the D1 locus. When all the loci are considered together, 2.4 percent of the total genomic diversity is attributable to the inter-population's variations. The total genomic diversity (H_T) among the sub-populations is high (0.495), whereas the H_S value, which determined the between individuals within populations, H_S is 0.483.

Table 3: Results of gene diversity analysis for individual loci and for all loci jointly considered in the study populations

Locus	H_T	H_S	G_{ST}
Alu FXIIIB	0.499456	0.4877	0.023602
Alu ACE	0.494445	0.4784	0.032356
Alu TPA25	0.499733	0.4848	0.029847
Alu mtNUC	0.499824	0.4962	0.007269
Alu PV92	0.481211	0.4788	0.005072
Alu APO	0.499003	0.4829	0.032219
Alu PLAT	0.496699	0.4935	0.006511
Alu D1	0.494275	0.4565	0.076444
Alu CD4	0.498016	0.4961	0.003901
All loci	0.495851	0.4839	0.024159

Genomic Affinities among Populations

The genomic affinities among four study populations are represented in Figure 1, using allele frequency data of nine loci by a standard NJ tree. This tree is divided into two clusters: Palliyan/Kattunaikkan and Malaikuravan/Malasar. It is seen that the affinities among the study populations do correlate well with their socio-cultural affiliation. Instead, populations that occupy closer geographical habitats show, by and large, closer genomic affinity. Pairwise genetic distances between the study populations were calculated from the allele frequencies using the DA distance measure in Table 4. The standard genetic distance ranged from 0.0133 to 0.0367 (between MR and KN). It was observed that Kattunaikkan is genetically more distant from other study populations.

Table 4: Matrices of pairwise distances among four south Indian tribal populations

	MK	*MR*	*PAL*	*KN*
MK	0.0000			
MR	0.0133	0.0000		
PAL	0.0217	0.0170	0.0000	
KN	0.0154	0.0367	0.0149	0.0000

To determine the genetic relationships of the present study, populations with other Indian tribal populations, the data of seven Alu indel loci (Alu mt NUC, Alu PV92, Alu FXIIIB, Alu APO, Alu ACE, Alu CD4 del and Alu PLAT) presented by Majumder et al. (1999); Mukherjee et al. (2000); Veeraju et al. (2008); Vishwanathan et al. (2003) and Dada et al. (2011) that are common with the present study were used. The NJ tree consisting of the 25 tribal Indian populations including the four study populations is presented in Figure 2. This tree shows that the tribes from central India clustered together while the four South Indian study populations stand apart genetically. In a similar way, the populations of Saharia, Koyadora and Kondareddi also get separated from other populations. In order to assess the global relationships of the study populations, the available Alu insertion data (ACE, FXIIIB, APO, PV92, PLAT) from Stoneking et al. (1997) was computed to construct the NJ tree (Fig.3). This NJ tree shows that the studied populations cluster with other Indian population namely, the Indian Christian cluster that include global populations.

Gene Flow among Populations

Harpending and Ward (1982) derived a regression of heterozygosity on genetic distance. This theory assumes a simple linear relationship between the heterozygosity of a population and the genetic distance of the population from the centroid (the overall mean allele frequency of the populations). If a population is receiving

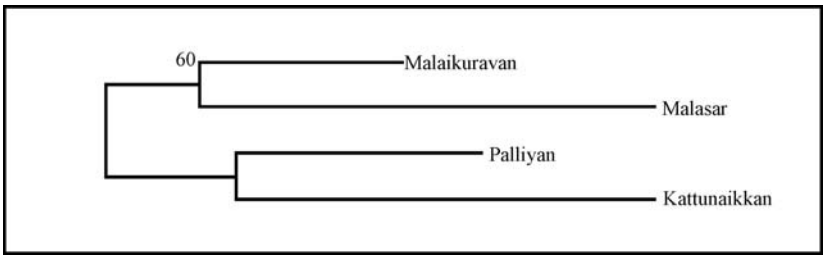

Fig. 1. Neighbor joining tree depicting genomic affinities among four tribal populations of South India

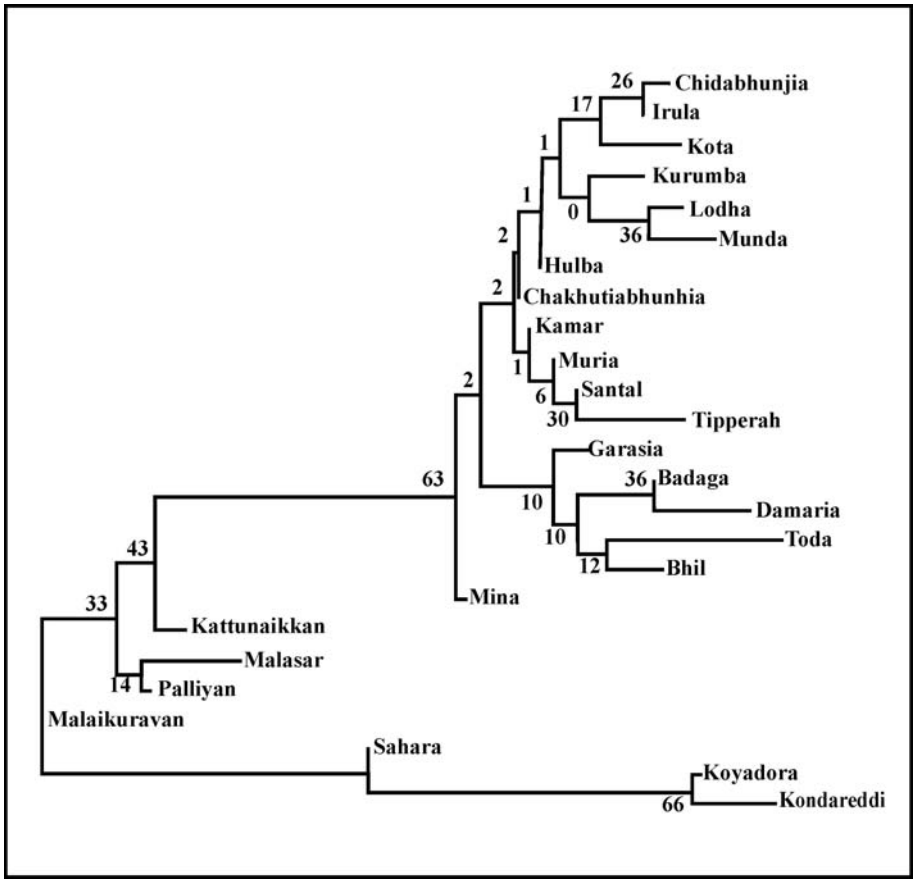

Fig. 2. Neighbor joining tree depicting genomic affinities of four tribal populations with other Indian tribal populations using seven Alu indels

genes from elsewhere at a higher than average rate, then the heterozygosity will be higher than predicted. If it is receiving genes at a lower than average rate, implying that the population is more isolated, the heterozygosity will be lower than predicted. To determine the relative amount of gene flow experienced by each population, a comparison of heterozygosity of each population against the genetic distance from centroid was performed. A centroid analysis has been

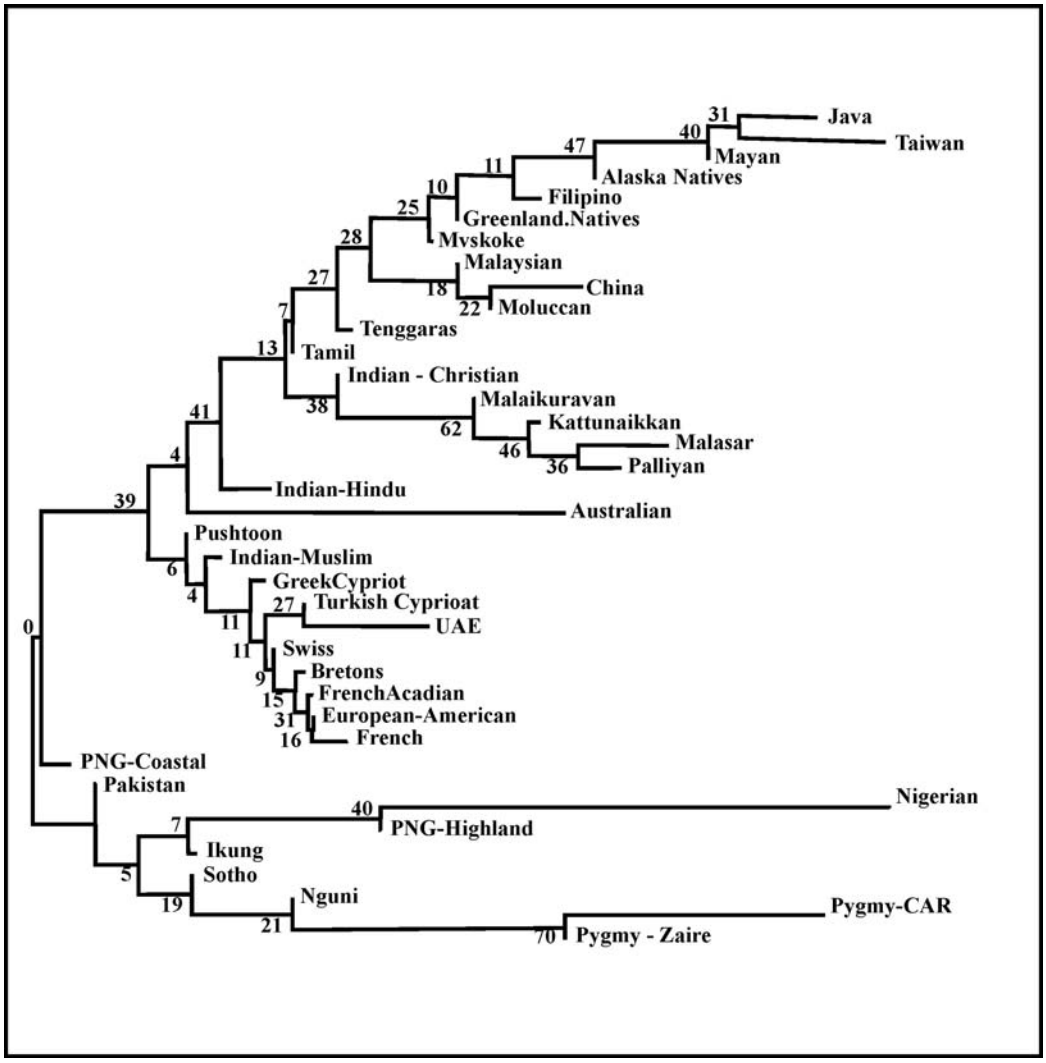

Fig. 3. Neighbor joining tree depicting genomic affinities of four tribal populations with world populations using five Alu indels

carried out by plotting heterozygosities of the four study populations and other Indian tribal populations by pooling the data with those (Majumder et al. 1999; Mukherjee et al. 2000; Vishwanathan et al. 2003; Veeraju et al. 2008; Dada et al. 2011) against the distance from gene frequency centroid is based on seven Alu insertion loci is presented in Figure 4. The analysis suggested that there could be moderate amount of gene flow between the set of populations under consideration.

DISCUSSION

Human specific insertion/deletion polymorphisms that have been used in this study are known to be selectively neutral in nature. Therefore, observed variations in the allele frequencies among populations are highly informative in assessing the genomic diversity of a population. The allele frequency distribution pattern of these populations was comparable with those observed in other Indian populations (Majum-

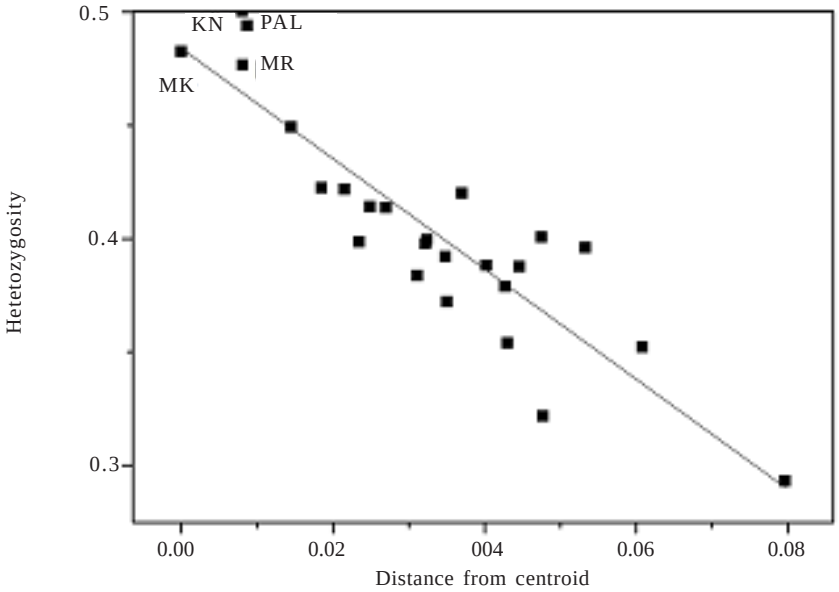

Fig. 4. Genomic affinities of the study population with other Indian tribal populations based on average heterozygosity Vs distance from centroid at seven Alu indel loci

der et al. 1999; Watkins et al. 2001; Vishwanathan et al. 2003; Veeraju et al. 2008; Kanthimathi et al. 2008; Yadav and Arora 2011; Dada et al. 2011; Kshatriya et al. 2011; Panjalia et al. 2012). Similarities of the allele frequency profile in all the four populations are due to the influence of their social structure and the strict endogamy practices.

Population relationships are shown by the topology of the NJ tree. The structure of the tree consists of two clusters: Malaikuravan/Malasar and Palliyan/Kattunaikkan. The comparison of gene frequencies in the above system shows variations, which implies that Palliyan and Kattunaikkan are more similar to each other compared to the other two tribal neighbors. The Malaikuravan and Malasar populations are genetically distinct. The genetic distance analysis indicates that the Palliyan and Kattunaikkan are closely related. The NJ tree analysis indicates that the study populations are grouped together and also cluster with other Indian Christian and world populations. This can be seen with the Kattunaikkan cluster and Mina, a tribal population of Rajasthan. The average heterozygosity values that range from 47.5 percent to 49.1 percent reflect the genetic heterogeneity in the study populations.

There is a significantly greater inter-individual variation within each study population than between the populations, hence the extent of population differentiation is rather low and the incident of average G_{ST} value for all markers in the four tribal groups 0.024 (2.4%). Earlier studies have reported G_{ST} values ranging from as low as two percent (Kshatriya et al. 2011) among eight western Indian tribal populations, to as high as eight point three percent among South Indian tribal populations (Vishwanathan et al. 2003).

The centroid analysis suggests that there is considerable amount of gene flow in the set of population under consideration. Further it can be explained that higher heterozygosity values and gene flow may be due to their small population sizes and closer proximity. The present study reveals that the four Dravidian tribal populations from South India are highly polymorphic, highly heterozygous in nature, with lower genomic differentiation and genetically distant from other Indian tribal and world populations. The probable explanation for above results could be their

small population sizes, strict endogamy practices and their geographical isolation for over a long period of time.

ACKNOWLEDGMENTS

The researchers are grateful to Dr. M.V. Usha Rani, Head of the Department of Environmental Science, Bharathiar University, Coimbatore, India, Mrs. S. P. Prasanna Vadhani and R. Prabhakaran for their help in collecting the samples and also the participants. The researchers also acknowledge the UGC, New Delhi for financial support.

REFERENCES

Bhasin MK 2009. Morphology to molecular anthropology: Castes and tribes of India. *Int J Hum Genet*, 9(3-4): 145-230.

Dada R, Saraswathy KN, Mettei KS, Mondal PR, Kaur H et al. 2011. Genetic sketch of the six population groups of Rajasthan: A study based on 12 autosomal loci. *Anthrop Science*, 119: 259-264.

Harpending HC, Ward R 1982. Chemical systematics and human evolution. In: MH Nitechi (Ed.): *Biochemical Aspects of Evolutionary Biology.* Chicago: University of Chicago Press, pp. 213-256.

Inanir A, Ceniklib A, Turalc E, Tekcanc A, Turalc S et al. 2014. Molecular analysis of genetic variation in angiotensin I-converting enzyme gene in Turkish athletes. *Int J Hum Genet*, 14(2): 101-105.

Kanthimathi A, Vijaya M, Ramesh A 2008. Genetic study of Dravidian caste of Tamil Nadu. *J Genet*, 87(2): 175-179.

Kshatriya GK, Aggarwal A, Khurana P, Italia YM 2011. Genomic congruence of Indo-European speaking tribes of western India with Dravidian-speaking populations of Southern India: A study of 20 autosomal DNA markers. *Ann Hum Biol*, 38(5): 583-591.

Majumder PP 1998. People of India: Biological diversity and affinities. *Evol Anthropol*, 6(3): 100-110.

Majumder PP, Roy B, Banerjee S, Chakraborty M, Dey B et al. 1999. Human - specific insertion /deletion polymorphisms in Indian populations and their possible evolutionary implications. *Eur J Hum Genet*, 7: 435-446.

Mukherjee N, Mitra M, Chakraborty M, Majumder PP 2000. Congruence of genomic and ethnolinguistic affinities among five tribal populations of Madhya Pradesh: *Ind J Genet*, 79: 41-46.

Nei M 1972. Genetic distance between populations. *Am Nat*, 106(949): 283-292.

Ota T 1993. *DISPAN: Genetic Distances and Phylogenetic Analysis.* Institute of Molecular Evolutionary Genetics. Pennsylvania State University, University Park, Pennsylvannia.

Panjaliya RK, Dogra V, Kumar P, Gupta S 2012. Human specific Alu insertion/deletion polymorphisms in various population groups of Jammu region. *Int J Hum Genet*, 12(4): 311-317.

Rakesh T, Lalji S, Kumarasamy T 2012. Complex genetic origin of Indian populations and its implications. *J. Biosci*, 37(5): 911-919.

Ray N 1973. *Nationalism in India.* Aligarh, India: Aligarh Muslim University.

Singh KS 1992. *People of India: An Introduction.* Anthropological Survey of India. Calcutta: Seagull Publication.

Singh KS 2010. *The Scheduled Tribes*. Oxford, UK: Oxford University Press.

Singh PK, Bourque G, Craig NL, Dubnau J T, Feschotte C et al. 2014. Mobile genetic elements and genome evolution. *Mobile DNA*, 5: 26

Stoneking M, Fontius JJ, Clifford SL, Soodyall H, Arcot S et al. 1997. Alu insertion/ deletion polymorphism and human evolution: Evidence for a larger population size in Africa. *Genome Res*, 7: 1061-1071.

Veerraju P, Demarche DA, Lakshmi N, Venkateswari rao T 2008. Insertion/deletion polymorphisms in Indian tribal populations. *Int J Hum Genet*, 8(1-2): 75-83.

Vishwanathan H, Deepa E, Usha Rani MV, Majumder PP 2003. Insertion/deletion polymorphisms in tribal populations of southern India and their possible evolutionary implications. *Human Biology*, 75(6): 873-887.

Watkins WS, Ricker CE, Bamshad MJ, Carroll ML, Nguyen SV, et al. 2001. Patterns of ancestral human diversity: An analysis of Alu insertion and restrict polymorphisms. *Am J Hum Genet*, 68: 738 – 752.

Yadav AB, Arora P 2011. Genomic diversity and affinities among eight endogamous groups of Haryana (India): A study on insertion/deletion polymorphisms. *Ann Hum Bio*, 38(1): 114-118.

Yeh FC, Yang RC 1999. *A Joint Project Development: POPGENE 1.32*. Centre for International Forestry Research. Canada: University of Alberta and Tim Boyle.