 IJHG 1(2): 129-132 (2001)

Insertion / Deletion DNA Polymorphisms in two South Indian Tribal Populations

P. Veerraju[1], T.V. Rao[1], N. Lakshmi[1], S. Reshmi[1], Badal Dey[2] and Partha P. Majumder[2]

1 Department of Human Genetics, Andhra University, Visakhapatnam 530 003, Andhra Pradesh, India
E-mail: pvraju@rediffmail.com
2 Anthropology and Human Genetics Unit,Indian Statistical Institute, Kolkata 700 035, West Bengal, India

KEY WORDS Alu insertion/deletion; mt. DNA; genome diversity; Koya Doras; Konda Reddis; average heterozygosity; genetic distance.

ABSTRACT DNA samples from 123 unrelated individuals belonging to 2 endogamous tribal populations, namely, the Koya Doras and Konda Reddis, of East Godavari District, Andhra Pradesh, South India were analysed for 8 human specific polymorphic insertion/deletion loci. Except Alu CD4, all the loci are polymorphic in both the populations. Most of the loci showed high levels of heterozygosity in both the populations. The average heterozygosity is higher in Konda Reddis than in Koya Doras. The results were compared with available Indian data on four other tribal populations, namely, Lodha, Munda, Santal and Tipperah.

INTRODUCTION

In recent times, polymorphic DNA markers are widely used to study the genomic diversity of Indian populations as most are selectively neutral, more ubiquitous and have higher heterozygosities than polymorphic protein and enzyme markers. Past studies on Indian ethnic groups have employed mitochondrial DNA markers (Mountain et al. 1995; Bamshad et al. 1996; Roychoudhury et al. 2000); Y chromosomal DNA polymorphisms (Thangaraj et al. 1998; Bhattacharyya et al. 1999) and VNTR and STR markers (Papiha et al. 1996; Mukherjee et al. 1999).

Indian society has been traditionally divided into a number of castes, tribes and religious communities. The tribals are accepted to be the autochthones. There are about 400 contemporary tribes (Majumder and Mukherjee 1993) constituting 7.5% of the total Indian population (Sirajuddin et al. 1994). The tribal people generally tend to be of shorter stature, more curly hair and darker pigmentation when compared to other Indians (Mountain et al. 1995).

In the present study eight human - specific insertion/deletion polymorphisms were studied in two tribal populations, namely, Koya Dora (KD) and Konda Reddi (KR). Of these, seven are Alu insertion/deletion markers, while the eighth marker (mt→NUC) pertains to a mitochondrial DNA segment, 540 bp in length, which got inserted into the human nuclear genome. It is also human – specific.

Though Alu sequences are found in most mammalian genomes, only some of them are human - spcecific (Batzer et al. 1994; Batzer et al. 1996). These human-specific Alu polymorphisms are very useful in the study of human population structure, because the ancestral states of these polymorphisms are known. The present study includes six human specific Alu polymorphisms (Alu D1, Alu ACE, Alu APO, Alu PV92, Alu TPA25, and Alu FX III B) whose ancestral state is the absence of the Alu element, while the ancestral state of Alu CD4 is the presence of the Alu element.

MATERIALS AND METHODS

Blood samples were collected, with appropriate consent, from 60 Konda Reddis (32 males and 28 females) and 63 Koya Doras (29 males and 34 females) belonging to nearby villages of Rampachodavaram of East Godavari district, Andhra Pradesh. All individuals were unrelated to the second – cousin level.

From each individual 5 ml of intravenous blood was collected in tubes containing EDTA. DNA was isolated using the salting out procedure (Miller et al. 1988). PCR reactions were carried out in a Perkin Elmer 2400 thermal cycler. Oligonucleotide primers used in PCR amplifications of the eight loci along with corresponding cycling conditions and reaction mixtures are given in table 1.

Gene counting method was used for estimating allele frequencies. Using these allele frequencies heterozygosities at individual loci were

Table 1: Oligonucleotide primers, reaction mixture and cycling conditions of the loci studied

Locus	*Primer sequences* *Cycling temperature protocol*	*Reaction mixture (10 µl)*	*References*
mt→NUC	5'- ACA AAG TCC AGG TTT CTA ACA G - 3' 5'- AGT CTT GCT TAT TAC AAT GAT GG – 3' 30 cycles x (94°C for 15s, 63°C for 30s, 72°C for 1min)	1.25 U Taq DNA polymerase 25 ng each primer, 100 ng genomic DNA	Zischler et al. 1995
Alu D1	5'- TGC TGA TGC CCA GGG TTA GTA AA – 3' 5'- TTT CTG CTA TGC TCT TCC CTC TC - 3' 30 cycles x (94°C for 1min, 66°C for 1min, 72°C for 1min)		
Alu ACE	5'- CTG GAG ACC ACT CCC ATC CTT TCT – 3' 5'- GAT GTG GCC ATC ACA TTC GTC AGA T – 3' 30 cycles x (94°C for 1min, 58°C for 1min, 72°C for 1min)		
Alu APO	5'- AAG TGC TGT AGG CCA TTT AGA TTA G –3' 5'- AGT CTT CGA TGA CAG CGT ATA CAG A – 3' 30 cycles x (94°C for 1min, 50°C for 1min, 72°C for 1min)	1.25 U Taq DNA polymerase 25 ng each primer 100 ng genomic DNA	Stoneking et al. 1997
Alu PV 92	5'- AAC TGG GAA AAT TTG AAG AGA AAG T – 3' 5'- TGA GTT CTC AAC TCC TGT GTG TTA G – 3' 30 cycles x (94°C for 1min, 54°C for 1min, 72°C for 1min)		
Alu TPA 25	5'- GTA AGA GTT CCG TAA CAG GAC AGC T – 3' 5'- CCC CAC CCT AGG AGA ACT TCT CTT T – 3' 30 cycles x (94°C for 1min, 58°C for 1min, 72°C for 1min)		
Alu FX III B	5'- TCA ACT CCA TGA GAT TTT CAG AAG T – 3' 5'- CTG GAA AAA ATG TAT TCA GGT GAG T – 3' 30 cycles x (94°C for 1min, 56°C for 1min, 72°C for 1min)		
ALU CD4	5'- AGG CCT TGT AGG GTT GGT CTG ATA – 3' 5'- TGC AGC TGC TGA GTG AAA GAA CTG – 3' 30 cycles x (94°C for 1min, 58°C for 1.5min, 72°C for 1min)	1.25 U Taq polymerase 25 ng each primer 100 ng genomic DNA	Edwards and Gibbs 1992

estimated. Gene diversity was estimated according to Nei (1973). Genetic distance was calculated as per Nei (1972). A dendrogram was constructed from the matrix of genetic distances following UPGMA (unweighted pair-group method) of Sneath and Sokal (1963). For gene diversity and cluster analyses, available data on other tribal populations of India (Majumder et al. 1999) were pooled with those generated in the present study. The other tribal populations include Lodha, Munda and Santal (Austro –Asiatic speaking tribals of West Bengal) and Tipperah (Tibeto- Burman speaking tribals of Tripura).

RESULTS

Allele Frequencies and Genomic Diversity Within Populations

The allele frequencies and heterozygosities for the insertion (+) /deletion (-) alleles for the eight loci studied in two populations, Koya Dora and Konda Reddi, are given in table 2. Except Alu CD4 all the loci are highly polymorphic. The deletion allele for Alu CD4 locus is completely absent in Konda Reddis. Most of the loci showed high levels of heterozygosity in both the populations. The average heterozygosity is higher in Konda Reddis than in Koya Doras.

Genomic Diversity Between Populations

Table 3 shows the results of gene diversity analysis in six tribal populations, two from Andhra Pradesh (the present study) three from West Bengal, the Lodha, Munda and Santal, and one from Tripura, Tipperah (Majumder et al. 1999). The total genomic diversity (H_T) among the sub populations is high. It is clearly evident from the table that most of the genomic diversity is due to the contribution of variation within populations (H_S) rather than genetic differences between populations. The diversity among the six populations is 4.8%.

Genomic Affinities among Six Tribal Populations

The affinities among six tribal populations (2 of the present study populations namely KD and KR and the 4 studied by Majumder et al. 1999, reconstructed employing the UPGMA method is shown in figure 1 using allele frequency data of all the eight loci. It is seen that

Table 2: Allele frequencies and heterozygosity at eight polymorphic loci in Koya Dora and Konda Reddi populations

Locus	*insertion/deletion* (+) (-)	*Koya Dora* n	*Konda Reddi* n
mt→NUC		55	56
	+	0.4273	0.4107
	–	0.5727	0.5893
heterozygosity		0.4894	0.4840
Alu D1		55	50
	+	0.4091	0.3200
	–	0.5909	0.6800
heterozygosity		0.4834	0.4352
Alu ACE		40	53
	+	0.7775	0.7170
	–	0.2250	0.2830
heterozygosity		0.3488	0.4058
Alu APO		58	54
	+	0.6724	0.6111
	–	0.3276	0.3889
heterozygosity		0.4406	0.4754
Alu PV92		55	53
	+	0.6545	0.6038
	–	0.3455	0.3962
heterozygosity		0.4522	0.4784
Alu TPA25		59	57
	+	0.6441	0.5965
	–	0.3559	0.4035
heterozygosity		0.4584	0.4814
Alu FXIIIB		55	50
	+	0.8000	0.5000
	–	0.2000	0.5000
heterozygosity		0.3200	0.5000
Alu CD4		59	54
	+	0.9746	1.0000
	–	0.0254	0.0000
heterozygosity		0.0496	0.0000
All loci heterozygosity		0.3803	0.4075

n = number of individuals studied

there is no clear clustering of ethnic groups by linguistic background, although the Tibeto-Burman speaking Tipperahs stand out as distinct from the Austro-Asiatic and Dravidian speaking tribal groups.

DISCUSSION

As new alleles are not generated at Alu insertion (+)/deletion (-) loci, and as there is no identified selection pressure on these loci, these loci have gained importance in the study of genetic structures of human populations. Several populations in India and other parts of the world, are highly polymorphic for these loci. Majumder et al. (1999) reported consistently high levels of average heterozygosity in 14 populations from India ranging from 0.351 to 0.449. The present study populations also exhibit high levels of heterozygosity. The KD has an overall heterozygosity of 0.3803 while KR has 0.4075. The extent of genomic differentiation (G_{ST}) among six tribal groups, Lodha, Munda, Santal (West Bengal) and Tipperah (Tripura) (Majumder et al. 1999) and KD and KR (present sudy) is higher than those observed by Stoneking et al. (1997) from other parts of the world except Africa. Majumder et al. (1999) observed that, consistent with the findings of classical markers, the Alu insertion/deletion markers also show high levels of genomic diversity in Indian populations. The results of the present study agree with this observation.

It is pertinent to note here that both the present study populations, KR and KD are linguistically Dravidian while three of the remaining four tribal populations are Austro-Asiatic (Santal, Munda and Lodha) and the fourth (Tipperah) is Tibeto Burman (Sino Indian). As is evident from the figure 1 the Tipperahs are genetically quite distant from the other populations. However, it is interesting to note here that

Table 3: Analysis of gene diversity in 6 tribal groups

Locus	H_T	H_S	D_{ST}	G_{ST}
mt → NUC	0.4890	0.4916	0.0064	0.013
Alu D1	0.4370	0.4334	0.0036	0.008
Alu ACE	0.4326	0.4068	0.0258	0.060
Alu APO	0.4758	0.4053	0.0705	0.148
Alu PV92	0.4768	0.4544	0.0224	0.047
Alu TPA 25	0.4866	0.4731	0.0135	0.028
Alu FXIIIB	0.3972	0.3692	0.0280	0.070
Alu CD4	0.0258	0.0256	0.0002	0.008
Mean	0.4037	0.3824	0.0213	0.048

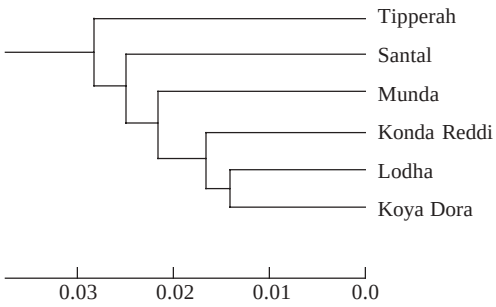

Fig. 1. Dendrogram of 6 Indian tribal populations based on diversity for eight polymorphic loci

though KR and KD are linguistically similar and geographically live in close geographical proximity, they do not form a cluster. It is observed that Lodhas and KDs though belonging to different linguistic groups and also inhabiting distant geographical locations, are genetically close and form a cluster. Thus, this analysis reveals that the Dravidian and Austro-Asiatic speaking tribals are genetically similar, while the Tibeto-Burman speaking Tipperah are genetically more distant from tribals of the other linguistic groups.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Department of Biotechnology (DBT), Government of India, New Delhi to P. Veerraju. Therefore we would like to thank the DBT for the same.

REFERENCES

Bamshad M, Fraley A, Crawford, MH, Cann RL, Busi BR, Naidu JM, Jorde LB 1996. mt DNA variation in caste populations of Andhra Pradesh, India. *Hum Biol*, **68**: 1-28.

Batzer MA, Stoneking M, Algeria-Hartman M et al 1994. African origin of human-specific polymorphic Alu insertions. *Proc Natl Acad Sci*, USA **91**: 12288-12292.

Batzer MA, Arcot SS, Phinney JW et al 1996. Genetic variation of recent Alu insertions in human populations. *J Mol Evol*, **42**: 22-29.

Bhattacharya NP, Basu P, Das M, Pramanik S, Banerjee R, Roy B, Roychoudhury S and Majumder PP 1999. Negligible male gene flow across ethnic boundaries in India, revealed by analysis of Y-chromosomal DNA polymorphisms. *Genome Research*, **9**: 711-719.

Edwards MC, Gibbs RA 1992. A human dimorphism resulting from the loss of an Alu. *Genomics*, **14**: 590-593.

Majumder PP, Mukherjee BN 1993. Genetic diversity and affinities among Indian populations. An overview. In: P.P. Majumder (Ed.). *Human Population Genetics*: A Centennial Tribute to J.B.S Haldane. New York: Plenum Press. pp. 255-275.

Majumder PP, Roy B, Banerjee S, Chakraborty M, Dey B, Mukherjee N, Roy M, Thakurta PG, Sil SK 1999. Human-Specific insertion deletion polymorphisms in Indian populations and their possible evolutionary implicatins. *Eur J Hum Gen*, **7**: 435-446.

Miller Sa, Dykes DD, Polesky HF 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*, **16**: 1215.

Mountain JL, Hebert JM, Bhattacharya S, Underhill PA, Ottolenghi C, Gadgil M, Cavalli-Sforza LL 1995. Demographic history of India and mt. DNA-sequence diversity. *Am J Hum Genet*, **56**: 979-992.

Mukherjee N, Majumder PP, Roy B, Roy M, Dey B, Chakraborty M, Banerjee S 1999. Variation of 4 short tandem repeat loci in 8 population groups of India. *Hum Biol*, **71**: 439-446.

Nei M 1972. Genetic distance between populatins. *Am Nat*, **106**: 283-292.

Nei M 1973. Analysis of gene diversity in subdivided populations. *Proc Nat Acad Sci*, **70**: 3321-3323.

Papiha SS, Mastana SS, Purandare CA, Jayasekara R, Chakraborty R 1996. Population genetic Study of Three VNTR loci (D2S44, D7SD22 and D12S11) in Five Ethnically Defined Populations of the Indian Subcontinent. *Hum Biol*, **68**: 819-835.

Roychoudhury S, Roy S, Dey B, Chakraborty M, Roy M, Roy B, Ramesh A, Prabhakaran N, Usha Rani MV, Viswanathan H, Mitra M, Sil SK, Majumder PP 2000. Fundamental genomic unity of ethnic India is revealed by analysis of mitochondrial DNA. *Curr Sci*, **79**: 1182-1192.

Sneath PHA, Sokal RR 1963. *Principles of Numerical Taxonomy.* San Fransisco: W.H. Freeman & Co.

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

Thangaraj K, Ramana GV, Singh L 1999. Y-chromosome and mitochondrial DNA polymorphisms in Indian populations. *Electriophoreses*, **20**: 1743-1747.

Zischler H, Geisert H, von Haeseler A, Paabo S 1995. A nuclear fossil of the mitochondrial D-loop and the origin of modern humans. *Nature*, **378**: 489-492.