

4

Genetic Relationships of Two Tibeto-Burman Speaking Tribal Groups of North East India with Two Other Indian Tribal Groups Speaking Different Languages

**C.S. CHAKRABARTI, M. ROY, N.K. SENGUPTA,
R. LALTHANTLUANGA, P.P. MAJUMDER**

INTRODUCTION

People of Indian subcontinent represent a great diversity of morphological, genetic, cultural and linguistic features (Majumder, 1998). To the characteristics of the indigenous populations of India, many exogenous features have been added due to immigrations that have taken place in historical times. As a result, the present Indian population comprises indigenous tribes with traditional life styles, along with many non-tribal groups. The number of tribal groups have been estimated to be about 450 (Singh, 1992). Some of the tribal groups are numerically large and found in various parts of India, while some groups are very small who inhabit restricted geographical territories. The tribal groups of India speak dialects that belong primarily to three major language families—Austro-Asiatic (a branch of Austric), Dravidian and Tibeto-

Burman (a branch of Sino-Tibetan). There are, however, a few tribal groups who speak dialects that belong to the Indo-European linguistic family. The Dravidian-speaking groups are primarily confined to the southern region of India, while the Tibeto-Burman speaking tribals are restricted to the northwestern region of India. In this study, we have attempted to estimate genomic diversities and relationships between two morphologically Mongoloid, Sino-Indian speaking tribal groups (Toto, Mizo) and have compared these with those of an Indo-European speaking tribal group (Tharu) inhabiting the sub-Himalayan region and a morphologically Proto-Australoid, Austro-Asiatic speaking tribal group (Ho).

The Tharu are a numerically large (about 30,000 individuals) tribe of Uttar Pradesh. Their origin is largely unknown. Their name may have been derived from 'Terai', a region in western Himalayan foothills, or from 'Thar' meaning 'forestation', or from 'Thar', a desert in western India. The Tharus claim to be descendants of the Rajputs, a dominant caste population of northern India, who are supposedly ancestors of many of the present-day Hinduised sections of the northern Indian ethnic groups. Some anthropologists (Majumdar, 1944) have claimed that they have a mixed ancestry. Because they have Mongoloid morphological features, it has been suggested that they may have been derived from admixture between the Proto-Australoid Indian tribals and some Mongoloid tribal group of Nepal. They are primarily agriculturists, although some are fishermen and some continue to be hunter-gatherers. They are divided into 5 main largely exogamous, subgroups. We have sampled individuals from two of these subgroups—Rana and Katharia.

The Totos are a very small, secluded primitive tribe comprising less than 1000 individuals confined to a particular village called Totopara (Sanyal, 1973). Totopara is situated 228—282 m. above sea level, in the district of Jalpaiguri in West Bengal, India, covering an area of approximately 2000 acres (Sanyal, 1973; Sarkar, 1993; Sinha and Pal, 1984). The Totos are morphologically, linguistically and culturally very different from the tribes and communities living around them, which include Koch, Rajbanshi, Mech, Garo, Tephy, Panikoch, Bhutia (Sanyal, 1973). Gates (1963) considered them as a distinct isolated tribe having Mongoloid features. The Toto society is male dominated, though women prefer husbands younger to them (Sinha, 1988). The Totos are an endogamous group having 13 exogamous clans (Sarkar, 1993). Inter-clan marriage is prohibited. Monogamy is the normally practiced (Sinha,

1988). Cross-cousin marriages, both maternal and paternal, are common (Debnath, 1982). Although there are many views regarding the origin of the Toto, the dominant view is that the Totos are a faction of the Bhutia community, having Tibetan ancestry.

The people who live in the north-eastern Indian State of Mizoram designate themselves as the Mizo ("the highlanders"). The State of Mizoram shares borders with Bangladesh and Myanmar. The origin of the Mizo tribals remains uncertain. But it is a fact that they entered from the direction of Myanmar in large numbers. The word 'Mizo' appears to be a blanket term and probably comprised many tribes, *e.g.*, Lushai, Punte, Poi, Roite, Darlong, Thado, Hmar, etc. Some of these tribes are no longer separately identifiable (*e.g.*, Darlong and Roite). However, the languages spoken by these groups, which belong to the Kuki-Chin sub-family, are dialectically and structurally so similar that it points to their belonging to a common ancestral stock (Goswami, 1979). They possess Mongoloid morphological features. Agriculture is the principal occupation of the Mizos. Most Mizos are followers of Christianity; a significant fraction also follow Buddhism. The total population size is about 450,000.

The Hos are an Austro-Asiatic speaking tribal group, predominantly inhabiting the eastern Indian region. Many anthropologists believe that the numerically large Austro-Asiatic tribals of this region were once a single tribe. Subsequently, in the course of migration to different areas, they assumed different names. The Austro-Asiatic speaking tribals are supposedly the most ancient inhabitants of India. The Hos have been living in isolation for a long time (Roy Chaudhury, 1958), in spite of the fact that they number over 400,000 individuals. Most are ancestor worshippers, although some have adopted Hinduism or Christianity. Their society comprises a number of exogamous clans. Cross-cousin marriages are not a taboo, if not one of the most preferred unions.

MATERIALS AND METHODS

Blood samples (5-10 ml in EDTA) were collected from individuals with prior their consent. All sampled individuals were unrelated at least to the first cousin level. The locations of sampling are indicated in Fig. 1. Blood samples were transported in ice to the laboratory of the Anthropology and Human Genetics Unit, Indian Statistical Institute. High molecular weight DNA was isolated from the blood samples by salting out procedure (Miller et al. 1988). Each DNA sample was

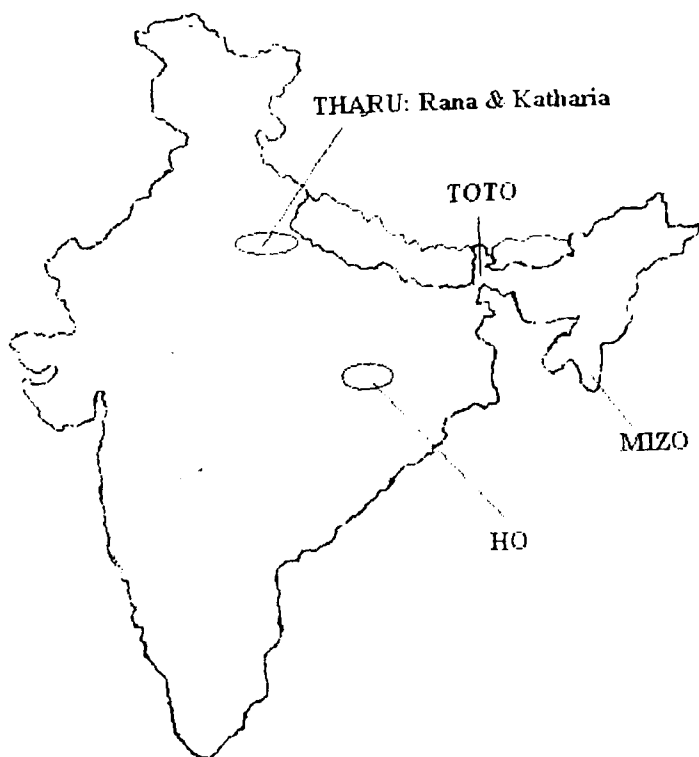


Figure 1. Map of India showing the locations of sampling of populations under study

analysed for polymorphisms at 25 loci, of which 8 were insertion/deletion polymorphisms (IDPs) and remaining 17 were RELPs. The names and GDB accession numbers or ALFRED (<http://alfred.med.yale.edu>) UID of the RSP loci are: ESRI (GDB:185229); NAT (GDB:187676); CYPIA (GDB: 9956062)-MspI; PSCR (GDB:182305); T2 (GDB: 196856); LPL (GDB: 285016); ALB(GDB: 178648); ALAD-MspI (GDB: 155925); ALAD-RsaI (GDB:155924); HBYB-Hinc II (GDB: 56084); HB 3ΨB-HincII; HB 5B—HinfI; Hox B4-MspI (UID: SI0001670); DRD2 (UID: SI000191L)- TaqIB, TaqID, Taq1A; ADH2-RsaI (UID: SI000002C). The names of the IDPs are given in Table 1. Primers and protocols

used for screening of the IDPs were as given in Majumder et al. (1999a) and Tishkoff et al. (1996), and those for RSPs were as given in Jorde et al. (1999a) and Tishkoff et al. (1996), and those for RSPs were as given in Jorde et al. (1995), Majumder et al. (1996b) and K. Kidd (personal communication). All DNA isolations and analyses have been performed in the laboratory of the Indian Statistical Institute, Kolkata.

Allele frequencies at each of these *biallelic loci* were estimated for each population by the maximum likelihood method. Chi-squared tests of significance between the observed genotype frequencies and those expected under Hardy-Weinberg equilibrium were performed. For the three sets of linked loci in and around the ALAD, DRD2 and Hb genes, maximum likelihood estimates of haplotype frequencies were estimated by the EM algorithm, using the HAPLOFREQ package (Majumder and Majumder, 2000). For analyses of genomic diversities, differentiation and distances, estimated haplotype frequencies at these three sets of linked loci were used. Observed gene diversities and the coefficient of gene differentiation G_{ST} , were estimated (Nei, 1973). Genetic distances between populations were estimated using the D_A , distance measure (Nei, 1983), An unrooted neighbour-joining tree (Saitou and Nei, 1987) was constructed to identify affinities among the tribal populations.

RESULTS AND DISCUSSION

Sample sizes and the + allele (Insertion allele for the Indel loci and presence of the restriction site for the RFLP loci) frequencies are given in Table 1. All the loci except *Alu* CD4 show high degrees of polymorphism across most populations. All populations at most loci show statistically non-significant differences of observed genotype frequencies and those expected under Hardy-Weinberg equilibrium Table 4.1.

However, all populations are significantly deviated from the Hardy-Weinberg equilibrium at the *Alu* D1 locus. At this locus, there is a significant excess of homozygotes over the expected. There are striking differences in allele frequencies at the PSCR locus; the Mizos are monomorphic at the locus. Among the 25 loci, there are three sets of linked loci in and around the ALAD, DRD2 and HB genes (gene clusters). Estimated haplotype frequencies for these sets of linked loci are given in Table 4.2.

Table: 4.1. Sample sizes, genotype and allele frequencies at 25 loci among tribal populations of India.

Locus	Katharia Tharu	Rana Tharu	Toto	Mizo	Ho
Alu MINuc	n	36	53	30	53
	+/+	9	20	15	5
	+/-	11	13	11	13
	-/-	16	20	4	9
p ± s.e.	0.403 ± .058	0.500 ± .048	0.683 ± 0.060	0.426 ± 0.067	0.453 ± 0.048
X ²	4.793*	13.755*	0.700	0.006	0.232
Alu PV92	n	36	53	30	29
	+/+	19	36	26	21
	+/-	12	15	4	8
	-/-	5	2	0	0
p ± s.e.	0.694 ± .548	0.821 ± .037	0.933 ± 0.032	0.862 ± 0.045	0.565 ± 0.048
X ²	1.657*	.077	0.153	0.742	0.0016
Alu FX3B	n	36	53	29	27
	+/+	22	35	24	10
	+/-	13	17	5	15
	-/-	1	1	0	2
p ± s.e.	0.729 ± 0.478	0.821 ± .037	0.914 ± 0.037	0.648 ± 0.065	0.802 ± 0.039
X ²	.323	.431	0.258	1.284	0.633

Contd.

Table 4.1 - (Contd . . .)

Locus	Katharia Tharu	Rana Tharu	Toto	Mizo	Ho	
Alu DI	n	36	53	30	28	
	+/+	12	19	8	7	
	+/-	6	9	4	2	
	-/-	18	25	18	19	
	p ± s.e.	0.417 ± .058	0.443 .048	0.333 ± 0.060	0.452 ± 0.049	23
Alu APO	X ²	15.546*	22.806*	14.700*	19.058*	
	n	36	53	30	29	
	+/+	25	26	23	29	
	+/-	11	25	5	5	
	-/-	0	2	2	2	
Alu ACE	p ± s.e.	0.847 ± 0.42	0.726 ± .043	0.850 ± 0.046	0.845 ± 0.048	0.849 ± 0.035
	X ²	1.171	1.848	3.600	3.400	1.675
	n	36	53	30	29	54
	+/+	12	21	9	10	30
	+/-	17	24	16	15	19
Alu ACE	-/-	7	8	5	4	5
	p ± s.e.	0.569 ± .058	0.629 ± .047	0.567 ± 0.064	0.603 ± 0.064	0.731 ± 0.043
	X ²	.049	.070	0.222	0.189	0.588

Locus	Katharia Tharu	Rana Tharu	Toto	Mizo	Ho
Alu CD4	n	36	54	30	26
	+/+	35	52	30	26
	+/-	1	2	0	0
	-/-	0	0	0	0
p ± s.e.	0.986 ± 0.14	0.981 ± .013	1.000	1.000	1.000
X ²	.007	.019	-	-	-
Alu PLAT	n	36	53	30	29
	+/+	24	32	0	10
+/-	11	20	2	14	25
-/-	1	1	28	5	3
p ± s.e.	0.819 ± .045	0.792 ± .039	0.033 ± 0.023	0.586 ± 0.065	0.713 ± 0.04
X ²	.038	1.148	0.036	0.001	0.928
ESR	n	37	54	30	29
	+/+	15	17	4	13
+/-	16	30	15	16	22
-/-	6	7	11	0	6
p ± s.e.	0.622 ± .056	0.593 .047	0.383 ± 0.063	0.724 ± 0.059	0.685 ± 0.04
X ²	.241	1.224	0.099	4.209*	0.167

Contd. . . .

Table 4.1 (Concid.)

Locus	Katharia Tharu	Rana Tharu	Toto	Mizo	Ho
NAT	n	37	54	30	26
	+/+	17	27	28	26
	+/-	20	24	2	0
	-/-	0	3	0	0
	p ± s.e.	0.730 ± 0.52	0.722 ± .043	0.967 ± 0.023	1.000
CYP1A	X ²	5.075*	.626	0.036	-
	n	37	52	30	26
	+/+	5	10	7	3
	+/-	15	28	12	10
	-/-	17	14	11	13
p ± s.e.	0.338 .055	0.462 ± .049	0.433 ± 0.064	0.308 ± 0.064	0.472 ± 0.048
PSCR	X ²	.326	.361	1.033	0.246
	n	37	54	30	26
	+/+	3	1	0	0
	+/-	9	17	3	0
	-/-	25	36	27	26
p ± s.e.	0.203 ± .047	0.176 ± .037	0.050 ± 0.028	1.000	0.264 ± 0.043
X ²	2.266	.397	0.083	-	0.046

Table 4.2. Haplotype frequencies at linked sets of loci in and around genes in tribal populations of India

Locus	Haplo- type	Katharia Tharu	Rana Tharu	Toto Ho	Mizo	Ho
ALAD	+/+	0.145	0.010	0.000	0.000	0.000
	+/-	0.313	0.188	0.033	0.111	0.071
	-/+	0.147	0.235	0.633	0.315	0.214
	-/-	0.395	0.567	0.333	0.574	0.714
DRD2	+/+/+	0.351	0.293	0.201	0.211	0.731
	+/+/-	0.079	0.204	0.164	0.163	0.000
	+/-/+	0.160	0.021	0.000	0.039	0.000
	+/-/-	0.035	0.011	0.000	0.000	0.000
	-/+/+	0.143	0.102	0.017	0.126	0.019
	-/+/-	0.052	0.014	0.000	0.000	0.000
	-/-/+	0.180	0.273	0.432	0.461	0.231
	-/-/-	0.000	0.082	0.185	0.000	0.019
Hb	+/+/+	0.216	0.084	0.300	0.130	0.040
	+/+/-	0.131	0.045	0.000	0.105	0.000
	+/-/+	0.041	0.152	0.000	0.026	0.000
	+/-/-	0.071	0.068	0.000	0.025	0.000
	-/+/+	0.115	0.220	0.050	0.132	0.180
	-/+/-	0.024	0.000	0.000	0.000	0.000
	-/-/+	0.239	0.346	0.650	0.569	0.760
	-/-/-	0.163	0.085	0.000	0.013	0.020

It is seen from the table that there is a large variation in estimated frequencies of haplotypes across populations. The differences in haplotype frequencies are not only significantly different at the 5% level among the 5 groups, but are also significantly different between the two Tharu subgroups. However, the two Tharu subgroups show similarities between themselves which is distinctive from the remaining three groups. The ALAD +/+ haplotype is only present among the two Tharu subgroups, but not in the other populations. Similarly, the DRD2 +/+ and the -/+/- haplotypes are not present among Totos, Mizos and

Hos, but are present among Rana and Katharia Tharus. The gene diversity for each of the study populations is given in Table 4.3. Gene diversity values range from 0.368 (Toto) to 0.467 (Katharia Tharu). It is not surprising that the Toto should have the lowest gene diversity, in view of its smallest population size among the study populations. It is also not surprising that the two subgroups of the Tharu have remarkably similar values of gene diversity.

Table 4.3. Gene diversities (on diagonal) and genetic distances between tribal populations of India

	<i>Katharia Tharu</i>	<i>Rana Tharu</i>	<i>Toto</i>	<i>Mizo</i>	<i>Ho</i>
Katharia Tharu	0.467 ± 0.041				
Rana Tharu	0.0164	0.462 ± 0.038			
Toto	0.0773	0.0517	0.368 ± 0.046		
Mizo	0.0213	0.0172	0.0536	0.448 ± 0.032	
Ho	0.0736	0.0634	0.0820	0.0576	0.376 ± 0.040

Genetic distances calculated on the basis of allele and haplotype frequencies at the 20 loci are provided in Table 3. A neighbour-joining tree (Fig. 2) was constructed. Bootstrapping was done using 1000 replications. It is seen from the reconstructed genomic affinities that the two subgroups of the Tharu are closer to each other. Even though most Tibeto-Burman speakers of northeast India are said to have originated from closely-related ancestral groups of southern China or southeast Asia, the Totos and the Mizos do not seem to be genetically close. The Tharu subgroups are closer to the Ho. The bootstrapped estimates of the confidence associated with these branches are quite high (Fig. 2). This confirms the earlier observations of anthropologists (Majumder, 1944) that the Tharu may be a mixed population. As a matter of fact, the speculation that the Tharu may have been derived from admixture between a Proto-Australoid tribal population of India and a Mongoloid tribal population of Nepal seems to be consistent with the findings of the present study. We have been unable to test this

hypothesis of admixture formally and to obtain statistical estimates of admixture proportions because comparable genotype data from Mongoloid tribal populations of Nepal are not currently available. It would be worth attempting in future.

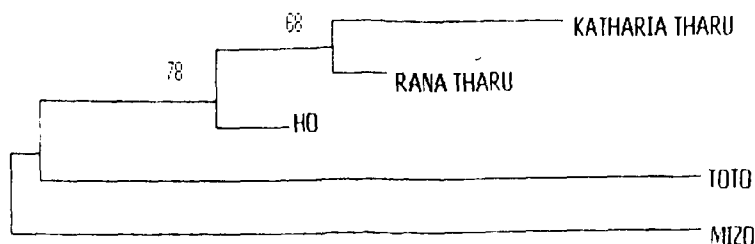


Figure: 2. Neighbour-joining tree depicting genomic relationships among tribal populations of India

ACKNOWLEDGMENTS

This study was supported by grants from the Department of Biotechnology, Government of India, and the Indian Statistical Institute. We are grateful to Dr. Ken Kidd and Dr. Lynn Jorde for sharing some laboratory protocols with us and to all members of the laboratory of the Anthropology & Human Genetics. Unit, Indian Statistical Institute, for help and co-operation throughout the course of this study.

REFERENCES

- Debnath, S.K. (1982). *Courtship and Ceremonial Aspect of marriage among Totos : a tribal isolated of Indo-Bhutan Border. Vanyajati, XXX : 49.*
- Gates, R.R. (1963). The Totos (Cited by Sanyal, C.C.1973. In: *Meches and the Totos : two sub-Himalayan tribes of the North Bengal.* Darjeeling: North Bengal University Press.
- Goswami, B.B. (1979). *Mizo Unrest: A Study of Politicization of Culture.* Jaipur: Aalekh Publishers.
- Jorde, L.B., Bamshad, M.J., Watkins, W.S., Zenger, R., Fralely, A.E., Krakowiak, P.A., Carpenter, K.D., Soodyal, H., Jenkins, T. & Rogers, A.R. (1995) Origins and Affinities of Modern Humans: a Comparison of Mitochondrial and Nuclear Genetic Data. *Am. J. Hum. Genet., 57,523-538.*
- Majumdar, D.N. (1944). *The Fortunes of Primitive Tribes,* Lucknow: Universal Publishers.
- Majumder, P.P. (1998). People of India : Biological Diversity and Affinities. *Evol. Anthropol., 6, 100-110.*

- Majumder, P.P., Roy, B., Banerjee, S., Chakraborty, M., Dey, B., Mukherjee, N., Roy, M., Guha Thakurta, P. & Sil, S.K. (1999a) Human-specific insertion/deletion Polymorphisms in Indian Populations and their possible Evolutionary implications. *Eur. J. Hum. Genet.* 7,435-446.
- Majumder, P.P., Roy, B., Balgir, R.S. & Dash, B.P. (1999b) Polymorphisms in the Beta-globin Gene Cluster in some Ethnic Populations of India and their Implications on Disease. In: *Molecular Intervention in Disease* (eds. S. Gupta & O.P. Sood), pp. 75-83. New Delhi: Ranbaxy Science Foundation.
- Majumder, P.P. & Majumder, P. (2000) *HAPLOFREQ: A Computer Program for Maximum Likelihood Estimation of Haplotype Frequencies in a Population from Genotype data on Unrelated individuals via the EM algorithm*. Technical Report, Calcutta: Indian Statistical Institute.
- Miller, S.A., Dykes, D.D. & Polsky, H.F. (1998) A Simple Salting out Procedure for extracting DNA from human nucleated cells. *Nucl. Acids Res.* 16, 1215.
- Nei, M. (1973) Analysis of Gene Diversity in Subdivided Populations. *Proc. Natl. Acad. Sci. USA* 70,3321-3323.
- Roy Chaudhury, P.C. (1958) *Bihar District Gazetteers : Singhbhum*. Patna: Secretariat Press.
- Saitou, N. & Nei, M. (1987) The Neighbour-Joining Method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406-425.
- Sanyal, C.C. (1973). *The Meches and the Totos : Two sub-Himalayan Tribes of North Bengal*. Darjeeling : North Bengal University Press.
- Sarkar, A. (1993). *Toto : Society and change* (A sub-Himalayan Tribe of West Bengal). Calcutta: Firma KLM Private Ltd.
- Singh, K.S. (1992). *People of India*. Delhi: Oxford India Press.
- Sinha, D. (1988). *Ethno-Zoological Studies of the Sub-Himalayan Tribal area of North Bengal with emphasis on Ecological and Bio-energetics of the Totos*. Ph.D.Thesis. Siliguri: North Bengal University.
- Sinha, D. & Pal, B.C. (1984). A study of Eco-sociological Change among and Underdeveloped Tribes : the Totos of North Bengal. *Bulletin of Cultural Research Institute*, Vol. XVI, No. 1-4.
- Tishkoff, S.A., Ruano, g., Kidd, J.R. & Kidd, K.K. (1996) Distribution and Frequency of a Polymorphic Alu insertion at the Plasminogen Activator Locus in Humans. *Hum. Genet.* 97, 759-764.