Insertion/Deletion Polymorphisms in Tribal Populations of Southern India and Their Possible Evolutionary Implications

H. VISHWANATHAN,[1] DEEPA EDWIN,[1] M.V. USHARANI,[1] AND P.P. MAJUMDER[2]

Abstract India has the unique distinction of having perhaps the largest diversities, both biological and cultural. The Nilgiri Hills of southern India, a home for several tribal pockets representing different genetic isolates, provides a genetic wealth to understand human evolution. We have analyzed eight widely distributed polymorphic insertion/deletion loci (*AluAPO*, *AluACE*, *AluD1*, *AluPLAT*, *AluPV92*, *AluFXIIIB*, *CD4* del and *mtNUC*) in 250 unrelated individuals from five tribal populations (Badaga, Irula, Kota, Kurumba, and Toda). All loci were highly polymorphic except the *CD4* del locus, at which the deletion allele was fixed in Kotas and Kurumbas. The levels of average heterozygosities were found to be high in all the populations. In most populations, they were also higher than those predicted by the island model of population structure. The gene diversity ($G_{ST} = 8.3\%$) was found to be higher than that in populations of most global regions with the exception of Africa. It is clear from the present study that drift effects could have accentuated the process of genetic differentiation of the tribal populations. The possibility of an early demographic expansion of modern humans within south India also cannot be ruled out.

Tremendous biological and cultural variability is manifested in the prehistoric and contemporary populations of India (Majumder 1998). Multiple waves of migration into India during prehistoric and historic times and the subsequent cultural differentiation resulting in strict rules governing mating practices are two of the major causes of the genomic diversity observed among contemporary ethnic populations of India. The contemporary people of India are culturally stratified as tribals and nontribals. The definition of tribe is somewhat ambiguous, but it often refers to the endogamous populations that are considered aboriginal inhabiting the Indian subcontinent before the immigration of pastoral nomads from western and central Asia (Cavalli-Sforza et al. 1994). These aboriginal populations represent the descendants of Paleolithic population expansions into South Asia. The source of these expansions is disputed, although it has been argued that Africa

[1]Division of Genetics, Department of Environmental Sciences, Bharathiar University, Coimbatore–641 046, India.

[2]Anthropology & Human Genetics Unit, Indian Statistical Institute, Kolkata–700 035, India.

Human Biology, December 2003, v. 75, no. 6, pp. 873–887.

KEY WORDS: INDIA, TRIBES, INSERTION/DELETION POLYMORPHISM, HUMAN GENOME DIVERSITY

may have made some genetic contribution to India, since some tribal populations in southern India possess what has been described as African (Negroid) physical characteristics (Maloney 1974; Saha et al. 1974; Roychoudhury 1982; Chandler 1988).

About 8.08% of the total Indian populations belong to one of the 461 tribal communities who speak about 750 dialects (Kosambi 1991), which can be classified into one of the following three language families: Austro-Asiatics, Dravidian, and Tibeto-Burman. Of these, the first two have been the major contributors to the development of Indian culture and society (Meenakshi 1995). The contemporary tribal populations are largely Dravidian or Austro-Asiatic. There is considerable debate about whether Dravidian languages owe their origin to Neolithic peoples of southern India or were brought into India (Misra 1992), but there is evidence that Dravidian speakers, who included settled agriculturists, predominated in both northern and southern India. South India is thus a nucleus of Dravidian culture. In spite of this linguistic homogeneity in southern India, the cultural barriers have probably been a major reason for genetic diversification among the people of this region. These populations are thus interesting and challenging for population genetics investigations due to their ethnic origins and interethnic admixture.

In view of the foregoing discussion it is of immense interest to consider patterns of genetic affinities among endogamous groups inhabiting smaller geographical regions, although most studies pertain to a very macro-level. Moreover, a vast majority of the earlier studies on genetic variability were based on classical genetic polymorphisms. The levels of polymorphism at loci that code for expressed proteins and enzymes are generally low because mutations at these loci are commonly deleterious and, therefore, are often strongly selected against. On the other hand, polymorphic DNA markers are widely used to study genomic diversity, since most are selectively neutral, more ubiquitous, and have higher heterozygosities than polymorphic protein and enzyme markers.

Thus, the increasing number and variety of molecular genetic markers offer additional opportunities for more detailed analysis of human evolution and of genetic diversity within and between human populations. The recent insertion of mobile elements of the *Alu* family in the human genome provides a distinct class of polymorphisms in the human genome (Deininger et al. 1999). The *Alu* insertion/deletion polymorphisms (indels) offer several advantages over other nuclear DNA polymorphisms because they are stable and unique markers that are identical by descent. They are highly informative in ascertaining relationships between individuals and populations because their ancestral states are known. They are easily typed and robust. Recently growing confidence in the usage of *Alu* indel markers has been well documented in human evolutionary studies (Stoneking et al. 1997; Majumder et al. 1999; Watkins et al. 2001; Batzer and Deininger 2002). We generated data on six human specific *Alu* polymorphisms, an insertion polymorphism of mtDNA in the human nuclear genome (Zischler et al. 1995), and a deletion of 256 base pairs (bp) of a 285-bp *Alu* element at the *CD4* locus (Edwards and Gibbs 1992). By and large, it is evident that only progressive accumu-

lation of data can help provide a better understanding of evolutionary history of the tribal populations and offer a better definition of the genetic landscape of the Indian subcontinent.

Methods

DNA Samples. We have studied a total of 500 chromosomes from unrelated individuals belonging to five endogamous Dravidian tribal populations of southern India. The tribal groups are confined to hilly tracts and valleys of the Nilgiri Hills of Tamil Nadu, in South India (Figure 1), located between 20 and 100 km away from each other. They are as follows: Badaga (n = 51), Irula (n = 50), Kota (n = 45), Kurumba (n = 54), and Toda (n = 50). These tribes have traditionally been hunter-gatherers and now practice shifting and/or settled cultivation. The population sizes of these groups vary from about 1300 (Toda) to about 150,000 (Badaga) individuals. Additional linguistic, historical, demographic, and genetic information about these populations has been described elsewhere (Thurston 1909; Saha et al. 1976; Breeks 1983; Singh 1994; Roychoudhury et al. 2000).

PCR Amplification. High molecular weight DNA was isolated from the blood samples by the salting-out procedure (Miller et al. 1988) and was suspended in 10m*M* Tris and 0.1m*M* EDTA for genotyping. All the polymorphic loci studied were genotyped by amplifying 200–300 ng of DNA in a standard 30-cycle three-step polymerase chain reaction (PCR). Appropriate annealing temperatures and additives were optimized for each system. The insertion/deletion polymorphisms studied are *AluAPO, AluACE, AluD1, AluPLAT, AluPV92, AluFXIIIB, CD4* del, and *mtNUC*. The protocols of the studied insertion/deletion polymorphisms have been reported previously (Stoneking et al. 1997; Majumder et al. 1999; Watkins et al. 2001). After PCR, 10 μL of the PCR product was subjected to electrophoresis in 1.5% 1 × TBE agarose gels with 0.5 μg of *Hae*III ø 174 DNA as a fragment-length marker. Ethidium bromide–stained gels were visualized by UV and were documented.

Data Analysis. Maximum likelihood estimates of allele frequencies and their standard errors were calculated at each locus separately for each population. Heterozygosities at individual loci and the average heterozygosity were calculated using the estimated allele frequencies for each population. Hardy-Weinberg equilibrium (HWE) was tested using the χ^2 goodness-of-fit test when the observed and expected numbers in all cells were large enough to meet the validity requirements; otherwise, an exact test was employed using a random permutation method with Bonferroni's correction for multiple comparisons using STATISTICA software. To assess the extent of gene differentiation among the population groups, Nei's (1973) measure of gene diversity was performed separately for each

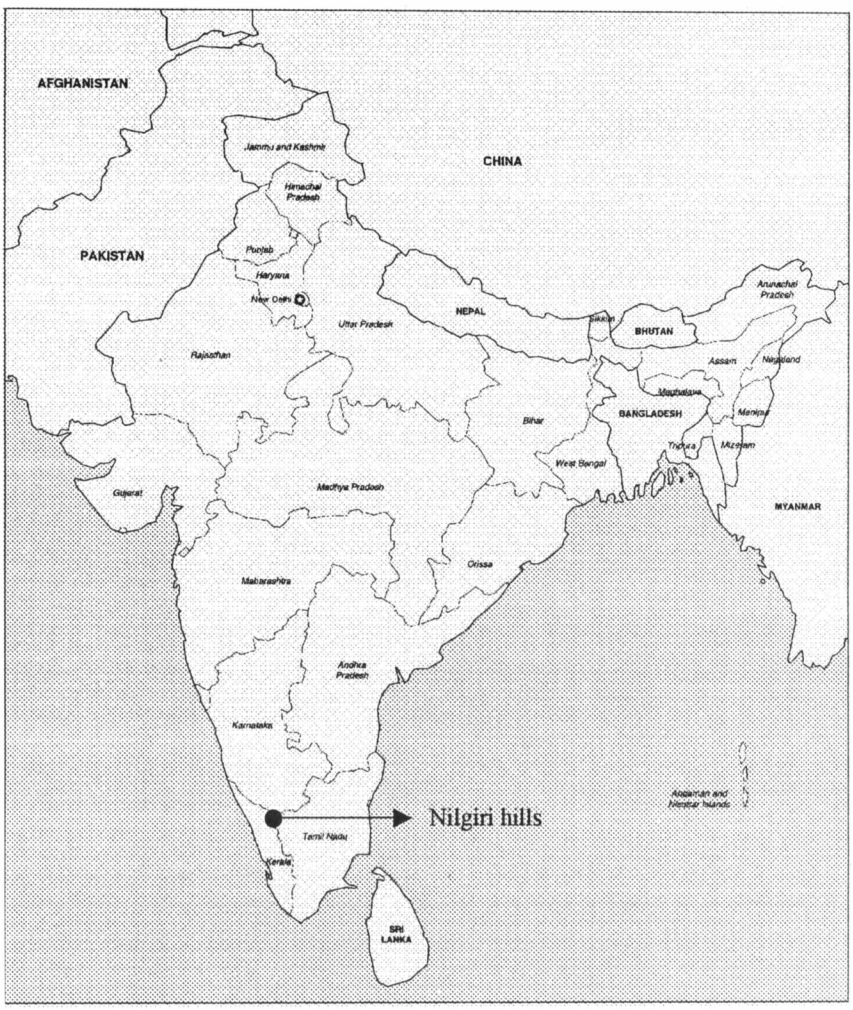

Figure 1. Map of India pinpointing the Nilgiri Hills in southern India.

locus and also for all loci considered jointly. To further test significant heterogeneity in the populations, which reflects differences in gene frequencies, the variation in genic populations was subjected to a contingency chi-square analysis (Workman and Niswander 1970). To assess genomic relationships among the populations, dendograms were constructed using the neighbor-joining (NJ) method. The relative amount of gene flow in each population was calculated by plotting the heterozygosity of each population against the distance of population from the centroid, as described by Harpending and Ward (1982).

Results

Genomic Diversity within Populations. The number of chromosomes examined and the allele frequencies (insertion [+] for *mtNUC* and *Alu FXIIIB, D1, APO, PLAT, ACE, PV92* loci; deletion [–] for *CD4* del locus) are presented in Table 1 separately for the five tribal populations. The (–) allele frequency for *CD4* del locus is presented because the deletion allele is human-specific. All the screened biallelic loci are highly polymorphic, with the exception of *AluAPO* in Todas, which was fixed for the absence of *Alu* element and *CD4* del, which exhibits low levels of polymorphism in all the populations. In fact, the deletion (–) allele is absent among the Kotas and Kurumbas. At most of the loci, all populations show statistically nonsignificant differences of observed genotype frequencies and of those expected under Hardy-Weinberg equilibrium. The heterozygosities at each locus and the average heterozygosities over all the eight loci for the study populations are given in Table 2. It is seen that all the populations show high levels of diversity with respect to most of the loci; the heterozygosity at the *CD4* del locus is low and consistently the minimum. The average heterozygosity for each locus was substantial, with most of the values approaching the theoretical maximum heterozygosity of 0.5 for a biallelic locus.

Genomic Diversity between Populations. To determine the amount of genetic differentiation among populations, G_{ST} values (a measure of the interpopulation variability) for each polymorphic locus were determined. The results of gene diversity are presented in Table 3, separately for each locus as well as for all loci pooled together. It is seen that except for the *CD4* del locus, the total genomic diversity (H_T) among the subpopulations is quite high. However, most of the genomic diversity is attributable to diversity between individuals within populations (H_S). The percentage of genomic diversity attributable to between populations relative to the total genomic diversity (G_{ST}) ranged from as high as 12.9% for *AluAPO* to as low as 1.9% for the *CD4* del locus. When all loci are jointly considered, 8.3% of the total genomic diversity is attributable to between populations. The tests of significance for heterogeneity of the populations showed significant values for 48 out of 80 comparisons (data not shown). These results

Table 1. Insertion/Deletion Frequencies in the Five Tribal Populations

	Badaga			*Irula*			*Kota*	
Locus	n^a	*p (+)*	*S.E.*	*n*	*p (+)*	*S.E.*	*n*	*p (+)*
AluACE	100	0.460	0.050	100	0.750	0.013	90	0.622
AluPLAT	98	0.551	0.050	100	0.550	0.050	88	0.659
AluPV92	94	0.436	0.051	98	0.449	0.050	90	0.300
AluAPO	102	0.784	0.041	100	0.570	0.050	90	0.767
AluFXIIIB	102	0.461	0.049	100	0.640	0.048	90	0.878
AluD1	100	0.230	0.042	100	0.600	0.049	90	0.589
CD4 (del)	100	0.020	0.014	100	0.040	0.020	90	0.000
mtNUC	96	0.625	0.049	94	0.574	0.051	90	0.567

a. n = Number of chromosomes.

	Kurumba			*Toda*		
S.E.	*n*	*p (+)*	*S.E.*	*n*	*p (+)*	*S.E.*
0.051	108	0.806	0.038	98	0.469	0.050
0.051	108	0.704	0.044	96	0.406	0.050
0.048	108	0.713	0.044	98	0.255	0.044
0.045	108	0.583	0.047	100	1.000	0.000
0.035	108	0.694	0.044	98	0.806	0.040
0.052	108	0.528	0.048	100	0.300	0.046
0.000	108	0.000	0.000	100	0.050	0.022
0.052	106	0.538	0.048	98	0.204	0.041

Table 2. Heterozygosities at Individual Loci and Average Heterozygosity Based on Eight Polymorphic Insertion/Deletion Loci

Locus	*Badaga*	*Irula*	*Kota*	*Kurumba*	*Toda*
AluACE	0.497	0.375	0.470	0.313	0.498
AluPLAT	0.495	0.495	0.449	0.416	0.484
AluPV92	0.491	0.495	0.420	0.409	0.385
AluAPO	0.338	0.490	0.357	0.486	0.000
AluFXIIIB	0.497	0.461	0.214	0.427	0.308
AluD1	0.260	0.160	0.067	0.129	0.206
CD4 (del)	0.039	0.076	0.000	0.000	0.095
mtNUC	0.469	0.490	0.491	0.497	0.325
All loci	0.386	0.380	0.309	0.335	0.288

Table 3. Gene Diversity Results at the Individual Locus and for All Loci Considered Jointly

Locus	H_T	H_S	G_{ST}
AluACE	0.471	0.431	0.085
AluPLAT	0.489	0.468	0.044
AluPV92	0.490	0.439	0.104
AluAPO	0.384	0.335	0.129
AluFXIIIB	0.423	0.381	0.098
AluD1	0.495	0.447	0.096
CD4 (del)	0.043	0.042	0.019
mtNUC	0.499	0.454	0.092
All loci	0.412	0.375	0.083

reflect substantial differences among populations with respect to their allele frequencies.

Genomic Affinities among Populations. In order to assess the relationship between the populations analyzed, genetic distance (D_A) was calculated and depicted in a neighbor-joining tree (Figure 2). Tree reconstruction methods depict population relationships as a series of bifurcations, which are commonly interpreted as population splits; however, it is important to realize that clusters of populations in such trees could arise from migration instead of from shared ancestry. A neighbor-joining (NJ) tree depicting the population relationships (Figure 2) divides into two clusters, which are: the Toda/Kota/Badaga and Irula/Kurumba.

To determine the genetic relationships of the tribal populations of the Nilgiri Hills with the tribal populations of India, the data of seven indel loci (*mtNUC, AluPV92, AluFXIIIB, AluAPO, AluACE, CD4* del, and *AluPLAT*) presented by

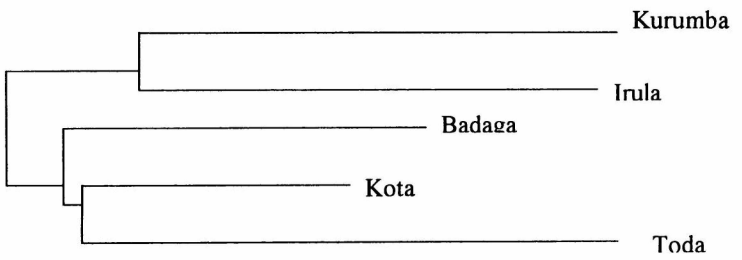

Figure 2. Unrooted NJ tree depicting genomic affinities among five south Indian tribal populations.

Majumder et al. (1999), Mukherjee et al. (2000), and Veerraju et al. (2001) that are common with the present study were used. The NJ tree of the 16 tribal Indian populations including the five tribal populations of the present study is presented in Figure 3. The tribal Indians with close geographic proximity to one another (central India) cluster together. It is seen that the Nilgiri Hills tribes—Toda, Kota, and Badaga—stand apart genetically, while the Irula cluster with the central Indians and Kurumba cluster with the Dravidian tribes of Andhra Pradesh. In order to compare the global relationships among populations, the available *Alu* insertion data (*ACE, FXIIIB, APO, PV92, PLAT*) from Stoneking et al. (1997) was computed. The overall pattern of the NJ tree showed that the Indian populations including the present study populations clustered together, with the exception of the Kurumba, between the Caucasoids and Mongoloids (Figure 4).

Gene Flow among Populations. To test whether in a group of incompletely isolated populations distributed over a geographical space (Wright's island model) the observed patterns of genomic diversities are the outcome of the process of drift and migration among the populations, or whether these patterns are generated by interactions with populations outside the set of populations under consideration, 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 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. The observed heterozygosities of the five populations against the distance from the gene frequency centroid are plotted along with the theoretical linear regression line (Figure 5). The analyses indicated

Figure 3. Unrooted NJ tree depicting genomic affinities among 16 Indian tribal populations based on seven insertion/deletion loci. Present study populations are in upper case; central Indian populations are in italics; Andhra Pradesh populations are underlined; northeast Indian populations are in bold.

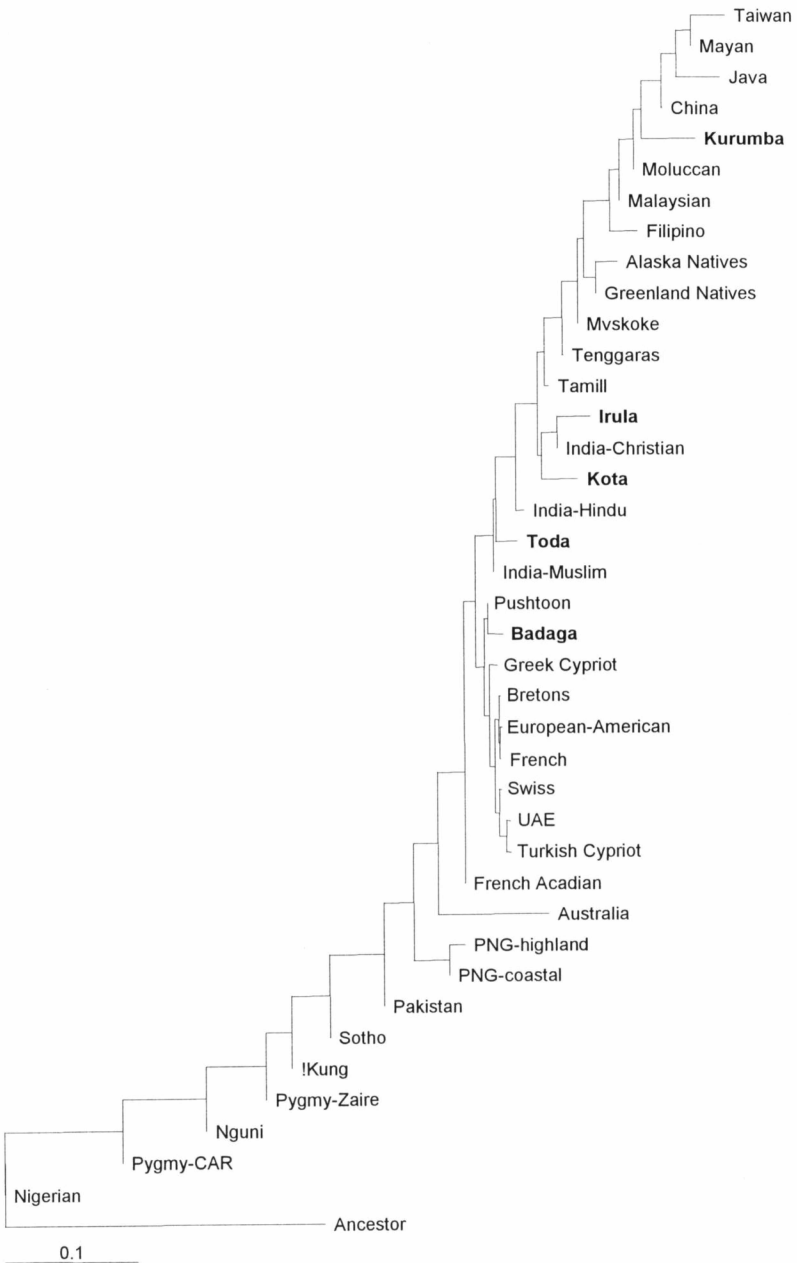

Figure 4. Neighbor-joining tree of global population relationships using five *Alu* indels. This tree is rooted where a hypothetical ancestral population—in which the frequency of the *Alu* element at each of the five loci is set to 0.0—attaches to the unrooted network.

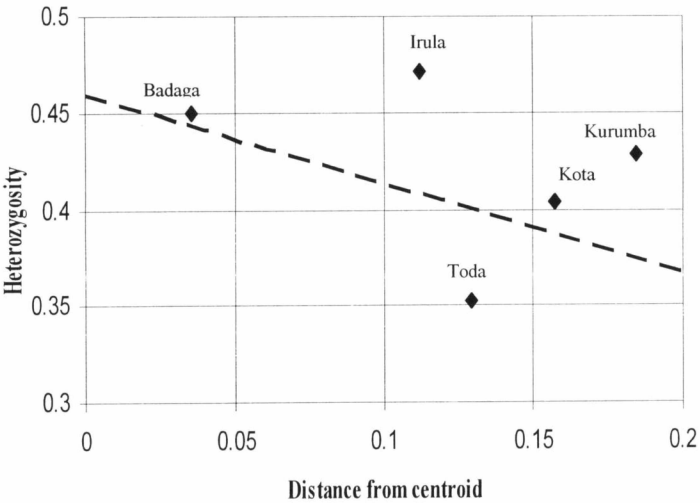

Figure 5. Plot of average heterozygosity Vs distance from centroid in the five tribal populations based on eight indel loci.

that the Toda populations have experienced less gene flow than predicted, while the Kotas, Badagas, Irulas, and Kurumbas have experienced considerably higher gene flow in terms of having higher than predicted heterozygosity.

Discussion

Though opinions differ, anthropologists, archeologists, and historians accept that the tribal populations are the original inhabitants of India. It is evident that south Indian tribals form a large and heterogeneous population, made up mostly of relic populations, which are of greater genetic interest (Cavalli Sforza et al. 1994). Thus the present investigation was more appropriate with the goal to analyze the extent of genomic variation at a number of insertion/deletion polymorphic loci from samples of diverse tribal populations of the Nilgiri Hills and to assess the relationships of these populations in terms of genomic affinity and diversity.

Four chi-square tests for goodness of fit to Hardy-Weinberg equilibrium were significant. When Bonferroni correction for multiple tests was computed, all the four significant deviations clustered on the *D1* locus. All the values that were significant are probably due to a deficiency in the observed heterozygotes, where, if they were normal statistical fluctuations, a similar number of departures from the expected number of heterozygotes for excess and deficiency would be expected. Another explanation is that the tribal groups reflect a true heterozygote deficiency, possibly because of inbreeding. Therefore, it may be concluded that the

deficiency in heterozygotes is real, possibly due to the reflection of a variable degree of inbreeding within the populations.

In the present study, the estimated levels of average heterozygosities are considerably high in all the populations. The heterozygosity levels are comparable to the average heterozygosity levels of other Indian populations (Majumder et al. 1999; Mukherjee et al. 2000; Veerraju et al. 2001). Interestingly, the average heterozygosity levels are higher than those of most other global populations studied with the exception of African populations (Stoneking et al. 1997; Novick et al. 1998). Thus, the DNA markers attest that the study groups exhibit high levels of genomic diversity.

The extent of genomic differentiation, G_{ST} estimates based on the eight polymorphic markers for the five tribal populations, is higher than that observed in all other parts of India (Majumder et al. 1999; Mukherjee et al. 2000; Veerraju et al. 2001). The G_{ST} values are smaller when the continental-level estimates were made on the basis of autosomal restriction site polymorphisms and short tandem repeat polymorphisms (Bowcock et al. 1991; Deka et al. 1995; Barbujani et al. 1997; Stoneking et al. 1997; Novick et al. 1998; Jorde et al. 2000). When the values of African populations were removed, however, the G_{ST} value reduced to 4.8%, which is lower than the observed value of the present study (8.3%). Recently, Watkins et al. (2001) reported 2.4% of G_{ST} estimates in the 12 Indian populations using *Alu* insertion polymorphisms, which was lower than half the estimated G_{ST} values of the present study. The comparison of significant heterogeneity also showed most, but not all, significant values. It is clear from the present investigation that drift effects have accentuated the process of genetic differentiations of the south Indian tribals.

Population relationships are shown by the topology of the NJ tree. The structure of the tree consists of two branches: Toda/Kota/Badaga and Kurumba/Irula (Figure 2). The comparison of gene frequencies in the above systems, which show variation, implies that Toda and Kota communities are more similar compared to the other tribal neighbors. The Toda and Kota are both distinctly different in physical features from the Australoid and Dravidian-speaking people of the area. Of interest, also, is the genetic distance between the Irula and Kurumba. These two groups are seen to be closely related. Indeed, the sampling of these two tribes comes from closer geographical proximity with very similar social and economic conditions. The Badagas, though clustering with the Toda and Kota, are distinctly different in physical features and in social and cultural practices. The possible reason might be due to recent admixture between these groups. Through the passage of time, genetic drift might have also played an important role in the genomic differentiation, in addition to other evolutionary forces like admixture and selection. The present analysis is in good concordance with the known population histories of the Nilgiri hill tribes and supports earlier reports (Chakravarthy and Mukherjee 1964; Saha et al. 1976; Ghosh et al. 1977) using classical markers. Thus, random genetic drift might have been due to their small population sizes and isolation in the hilly terrain. The significant genetic heterogeneity and the

clear clustering of these populations into groups of two in such close geographical proximity attest to the sensitivity of the demographic and genetic approaches in unraveling human history.

The genetic relationships between the present study populations and the other Indian tribal populations show that the study groups stand apart genetically, while the Irulas and Kurumbas have a remote genetic relationship with the tribes in other parts of India (Figure 3). This observation could indicate a common ancestry of Irulas and Kurumbas with other Indian tribal populations, an observation that is in good agreement with Labie et al. (1989), who postulated a unicentric origin of the tribal populations of India on the basis of their β-globin haplotypes. In the NJ tree, the other three groups were characterized with long branches, consistent with genetic drift occurring in these populations. Further, the relationship with the global populations using *Alu* insertion markers, which is shown in the NJ tree in Figure 4, is consistent with that reported by Majumder et al. (1999), which states that the Indian populations are genetically in between the Caucasoids and Mongoloids.

The centroid analysis (Figure 5) also shows that there has been a considerable amount of gene flow between the set of populations under consideration. In spite of considerable gene flow as inferred from the centroid analysis, the extent of gene differentiation among south Indian populations continues to be very high. Since this analysis does not permit timing of the period during which gene flow might have occurred between these south Indian populations, it is difficult to offer a clear interpretation of this finding. A probable explanation is that gene flow occurred prior to the subdivision of these south Indian populations into distinct endogamous units.

Further, it was found that all the populations except Todas showed considerably higher levels of heterozygosity than that predicted by the Harpending-Ward (Harpending and Ward 1982) gene flow model. If this pattern of high heterozygosities were simply due to higher levels of gene flow, then one would have expected that south Indian populations would be genetically less differentiated. However, it was found that the coefficient of gene differentiation among south Indian populations is higher than that among populations inhabiting all other regions of the world, except Africa (Stoneking et al. 1997; Majumder et al. 1999; Watkins et al. 2001). Therefore, it is reasonable to speculate that population bottlenecks and genetic drift have been important factors in generating high genetic differentiation in these populations. Since the joint observation of higher-than-predicted heterozygosities and a high level of genetic differentiation have earlier been accepted as hallmarks of population expansion (Stoneking et al. 1997; Batzer et al. 1996; Majumder et al. 1999), the possibility of an early demographic expansion of modern humans within south India cannot be ruled out.

Acknowledgments We are grateful to the Department of Biotechnology, Government of India, for financial support for this study. We also thank Richard Cordaux, MPI, Ger-

many, and the members of the laboratory of the Anthropology and Human Genetics Unit of the Indian Statistical Institute, Kolkata, for their help at various stages of this work. D.E. was supported by CSIR, India.

Received 15 November 2002; revision received 23 June 2003.

Literature Cited

Barbujani, G., A. Magagni, E. Minch et al. 1997. An apportionment of human DNA diversity. *Proc. Natl. Acad. Sci.* 94:4516–4519.

Batzer, M.A., S.S. Arcot, J.W. Phinney et al. 1996. Genetic variation of recent *Alu* insertions in human populations. *J. Mol. Evol.* 42:22–29.

Batzer, M.A., and P.L. Deininger. 2002. *Alu* repeats and human genome diversity. *Nature Rev. Genet.* 3:370–379.

Bowcock, A.M., J.M. Hebert, J.L. Mountain et al. 1991. Study of an additional 58 DNA markers in five human populations from four continents. *Gene Geography* 5:151–173.

Breeks, J.W. 1983. *Primitive Tribes and Monuments of the Nilgiris.* Delhi, India: Cultural Publishing House.

Cavalli-Sforza, L.L., P. Menozzi, and A. Piazza. 1994. *History and Geography of Human Genes.* Princeton, NJ: Princeton University Press.

Chakravarthy, M.R., and D.P. Mukherjee. 1964. Dermatoglyphic affinities of the tribes and castes of Nilgiri Hills (India). *Z. Morph. Anthrop.* 55:335–356.

Chandler, W.B. 1988. The jewel in the lotus: The Ethiopian presence in the Indus valley civilization. In *African Presence in Early Asia,* I.V. Sertma and R. Rashidi, eds. New Brunswick, NJ: Transaction, 80–105.

Deininger, P.L., S.T. Sherry, G. Risch et al. 1999. Interspersed repeat insertion polymorphisms for studies of human molecular anthropology. In *Genomic Diversity: Applications in Human Population Genetics.* New York, NY: Kluwer Academic Publishers.

Deka, R., L. Jin, M.D. Shriver et al. 1995. Population genetics of dinucleotide $(dC\text{-}dA)_n$. $(dG\text{-}dT)_n$ polymorphisms in world populations. *Am. J. Hum. Genet.* 56:461–474.

Edwards, M.C., and R.A. Gibbs. 1992. A human dimorphism resulting from the loss of an *Alu*. *Genomics* 14:590–593.

Ghosh, A.K., R.L. Kirk, S.R. Joshi et al. 1977. A population genetic study of the Kota in the Nilgiri Hills, South India. *Hum. Hered.* 27:225–241.

Harpending, H.C., and R.H. Ward. 1982. Chemical systematics and human populations. In *Biochemical Aspects of Evolutionary Biology,* M.H. Nitechi, ed. Chicago, IL: University of Chicago Press.

Jorde, L.B., W.S. Watkins, M.J. Bamshad et al. 2000. The distribution of human genetic diversity: A comparison of mitochondrial, autosomal and Y-chromosome data. *Am. J. Hum Genet.* 66:979–988.

Kosambi D.D. 1991. *The Culture and Civilization of Ancient India in Historical Outline.* New Delhi, India: Vikas.

Labie, D., R. Srinivas, O. Dunda et al. 1989. Haplotypes in tribal Indians bearing sickle gene: Evidence for the unicentric origin of the beta^S mutation and the unicentric origin of tribal populations of India. *Hum. Biol.* 61:479–491.

Majumder, P.P. 1998. People of India: Biological diversity and affinities. *Evol. Anthropol.* 6:100–110.

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

Maloney, C. 1974. *The Races in Peoples of South Asia.* New York, NY: Holt, Rinehart & Winston.

Meenakshi, K. 1995. *Recent Perspectives of Early Indian History,* R. Thapare, ed. Mumbai, India: Popular Prakashan, 1–52.

Miller, S.A., D.D. Dykes, and H.F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:1215.

Misra, S. 1992. The age of the Acheulian in India: New evidence. *Curr. Anthropol.* 33:325–327.

Mukherjee, N., M. Mitra, M. Chakraborty et al. 2000. Congruence of genomic and ethnolinguistic affinities among five tribal populations of Madhya Pradesh, India. *J. Genet.* 79:41–46.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70(12):3321–3323.

Novick, G.E., C.C. Novick, J. Yunis et al. 1998. Polymorphic *Alu* Insertions and the Asian origin of Native American populations. *Hum. Biol.* 70:23–39.

Roychoudhury, A.K. 1982. Genetic relationships of Indian populations. In *Human Genetics and Adaptation*, K.C. Malhotra and A. Basu, eds. New York, NY: Plenum Press, 147–174.

Roychoudhury, S., S. Roy, B. Dey et al. 2000. Fundamental genomic unity of ethnic India revealed by analysis of mitochondrial DNA. *Curr. Sci.* 79(9):1182–1192.

Saha, N. et al. 1974. Genetic studies among the Kadars of Kerala. *Hum. Hered.* 24:198–218.

Saha, N., R.L. Kirk, S. Shanbhag et al. 1976. Population genetic studies in Kerala and Nilgiris (south-west India) *Hum. Hered.* 26:175–197.

Singh, K.S. 1994. *The People of India: The Scheduled Tribes.* Oxford, UK: Oxford University Press.

Stoneking, M., J.J. Fontius, S.L. Clifford et al. 1997. *Alu* insertion polymorphisms and human evolutions: Evidence for a larger population size in Africa. *Genome Res.* 7:1061–1071.

Thurston, E. 1909. *Castes and Tribes of Southern India.* Vol. 4. Madras, India: Government Press.

Veerraju, P., T. V. Rao, N. Lakshmi et al. 2001. Insertion/Deletion DNA polymorphisms in two south Indian tribal populations. *Int. J. Hum. Genet.* 1(2):129–132.

Watkins, W.S., C.E. Ricker, M.J. Bamshad et al. 2001. Patterns of ancestral human diversity: An analysis of *Alu*-insertion and restriction-site polymorphisms. *Am. J. Hum. Genet.* 68:738–752.

Workman, P.L., and J.D. Niswander. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. *Am. J. Hum. Genet.* 22:24–49.

Zischler, H., H. Geisert, A. Von Haeseler et al. 1995. A nuclear 'fossil' of the mitochondrial D-loop and the origin of modern humans. *Nature* 378:489–492.