# Antiquity, geographic contiguity and genetic affinity among Tibeto-Burman populations of India: A microsatellite study

S. KRITHIKA[1], R. TRIVEDI[2], V. K. KASHYAP[3], P. BHARATI[1], & T. S. VASULU[1]

[1]*Biological Anthropology Unit, Indian Statistical Institute, Kolkata, India,* [2]*National DNA Analysis Center, Central Forensic Science Laboratory, Kolkata, India,* and [3]*National Institute of Biologicals, Noida, India*

## Abstract

*Background:* The Tibeto-Burman (TB) populations are one of the four major linguistic population groups of India. They are considered belonging to different stocks and show wide variation in culture and language; however, their genetic relationship, antiquity and migration history among the regional populations has been little investigated. Molecular genetic studies are expected to clearly show the antiquity and genetic diversity of these populations.

*Aim:* This paper seeks to understand the extent and magnitude of genetic affinities and diversities among 14 TB populations (12 Indian and two global groups), investigate the findings based on classical genetic markers and verify the historical accounts of their migration and genetic history based on 12 microsatellite markers.

*Subjects and methods:* The allele frequency data for 12 STR loci of 13 Asian (Tibeto-Burman) populations were obtained from the literature and the Adi Pasi data was obtained by microsatellite typing of their blood samples. The 12 loci studied are D5S818, FGA, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, vWa, TPOX, D18S51. Three different distance measures, two phylogenetic trees and PCA plot have been employed to understand the genetic relationship of the studied populations.

*Results:* Average heterozygosity values range from 68 to 79% and the average $G_{ST}$ value is 4.7%. The dendrogram, based on the $D_A$ distance, shows the clustering of populations based on their diversities and geographical contiguity; the Mizoram and Arunachal Pradesh populations especially cluster together, populations from Sikkim form a separate subcluster and Manipur populations along with the Garo of West Bengal separate out from the other clusters. The Harpending and Ward regression model shows isolated populations positioned below the regression line and others, who experience external gene flow, placed above the line. The results support folklore migration accounts of their possible antiquity with the Tibetan and southern Chinese populations.

*Conclusions:* Overall, geographic contiguity, punctuated by isolating barriers, is a major influencing factor of genetic affinity among the TB population; contiguous populations within a region

show greater genetic relationship than with distant TB populations over a wide geographical area. The results of the microsatellite study also support the history of diverse routes of migration of these populations.

**Keywords:** *Tibeto-Burman populations, microsatellite markers, northeast India*

## Introduction

Tibeto-Burman speaking tribes are one of the four major linguistic groups in India. They are confined to the north and northeastern parts (Grierson 1903–1909; Ruhlen 1991) and constitute less than 2% of the country's total population (Malhotra and Vasulu 1993). Compared to Austro-Asiatic, Dravidian and Indo-European speaking tribes, Tibeto-Burman speakers are supposed to be recent migrants to the Indian subcontinent (Gadgil et al. 1997). They are assumed to have arrived at different periods from eastern, southeastern (Rapson 1955; Dani 1960) and central Asian regions. However, further details of their origin, route of migration, settlement in different parts of the country and the genetic affinity among these populations is not clearly known. This plethora of diverse stocks is evident from the variation in culture, language, subsistence economy and demography and from population structure variables like size, marriage patterns, endogamy etc. (Elwin 1959; Majumdar 1980; Singh 1998).

Tibeto-Burman populations constitute a significant component of the biological diversity of the peopling of India and provide some distinct advantages for the investigation of population genetics and micro-evolutionary processes. For example, they show wide variation in population structure in different regions; from hunting-gathering to civilized life styles, the biological consequences of which are of much importance. Some of these tribes that live in relative isolation, undisturbed by external influences and following a hunting-gathering life style, provide advantages for genetic studies (WHO 1964), which could facilitate the discovery of some rare alleles or traits. In contrast, some other tribes, living in the plains with the castes and representing a population structure of urban life style, show wide genetic diversity as a result of cultural and genetic admixture (Kumar et al. 2004). Studies of the diverse population structure of these populations are of great significance in understanding the peopling of India, affinity and diversity among the regional populations, and their relationship with southeast Asian populations. However, there are very few such studies of the extent and magnitude of this genetic diversity and/or genetic affinity (Das et al. 1980; Roychoudhury 1981; Walter et al. 1986; Deka et al. 1988; Roychoudhury 1992; Kumar et al. 2004).

There are different hypotheses concerning the antiquity, origin and migration of Tibeto-Burman speaking tribes. Folklore tradition and cultural materials indicate their antiquity and possible origin from Tibet. Some migrated to north and northeast India, and others are presumed to have traveled to Burma and later settled in northeastern parts of the country. If the above hypotheses are true, it is expected that geographically proximate populations should show greater genetic affinity than distantly located populations. Therefore, based on regional distribution, some northeastern populations in Manipur, Tripura are expected to show close genetic relatedness to southeastern populations and populations in Sikkim, Ladakah region are expected to show close affinity with southern Chinese and Tibetan populations. However, some populations in Arunachal Pradesh, Mizoram, close to Burma, and southern China, possibly came from either Burma or southern China or from both directions. As such they might show affinity with both populations across the border. The biological validity of the above hypothesis has rarely been

investigated using classical genetic markers, except for a few studies based on some specific regional populations. A few studies based on classical genetic markers among the TB populations in the eastern region have observed genetic affinity among regional populations of close geographical proximity, no matter whether they originated from the same tribal group or same linguistic family in the past (Bhasin et al. 1992; Roychoudhury 1992; Papiha et al. 1996; Majumder 1998a, 1998b).

In recent years several populations have been screened for molecular genetic markers like microsatellites, mtDNA and Y-chromosome markers, primarily to develop a DNA data bank which helps investigation of past genetic history, origin and migration, genetic affinity and diversity of regional populations. It also improves understanding of the genetics of diseases prevalent among these populations and clarifies findings based on classical genetic markers and historical accounts (Cann et al. 1987; Bhattacharya et al. 1999; Underhill et al. 2001; Kivisild et al. 2002; Basu et al. 2003). Until now, very few molecular genetic studies among Indian Tibeto-Burman populations have been carried out and these were sporadic and limited to only some regional populations (Dutta et al. 2001; Chakrabarti et al. 2002; Dutta et al. 2002; Cordaux et al. 2003; Kashyap et al. 2004a).

Our present study considers all available microsatellite data on various Tibeto-Burman populations and utilizes them to understand different aspects of these populations from a genetic viewpoint. We have studied 12 microsatellite (STR) markers among a set of populations of the Tibeto-Burman linguistic family from north and northeastern parts of the country along with two global Asian (Tibeto-Burman) populations: (a) to investigate the extent and magnitude of genetic affinities and diversities among regional and global Tibeto-Burman populations, (b) to infer the biological validity of the diverse origin and migration of different populations using STR markers as these markers are supposed to reflect their recent genetic history and (c) to verify the relative influence of geographic proximity on genetic affinity among regional populations (findings based on classical genetic markers).

## Materials and methods

### Populations

The study comprises 14 populations that include 13 Tibeto-Burman speaking Asian or Asian derived tribal populations and one population of western Eurasian origin. Out of 14 Tibeto-Burman populations, 12 are from different parts of India and two are from China and Malaysia respectively. The Indian populations include Hmar, Mara, Lai and Lusei of Mizoram; Nepali, Bhutia, Lepcha of Sikkim; Kuki and Naga of Manipur; Garo of West Bengal and Adi Pasi of Arunachal Pradesh. The Muslims of Manipur (MM) were included as an outlier in the study. The two global populations are Han Chinese from China and Javanese from Malaysia. A map of northeastern India showing the distribution of the studied populations is depicted in Figure 1. Details of the studied populations, their geographical locations, occupation and linguistic affiliations are given in Table I.

### DNA isolation and microsatellite typing

Blood samples were collected with prior informed consent from 203 healthy voluntary donors belonging to the Adi Pasi subtribe from seven villages located at low altitude in the

1- Adi Pasi, 2- Kuki, 3- Naga, 4- Muslims, 5- Hmar, 6- Mara, 7- Lai,
8- Lusei, 9- Nepali, 10- Bhutia, 11-Lepcha, 12- Garo, 13- Chinese

Figure 1. Map of northeastern India showing the geographic distributions of the studied populations.

Table I. Sample size, geographic locations, occupational, ethnic, and linguistic backgrounds of the studied populations.

| Population | Sample size | Location | Subsistence pattern | Linguistic affiliation** (branch) |
|---|---|---|---|---|
| Adi Pasi | 203 | Arunachal | Hunting gathering | North-Assam |
| Hmar | 80 | Mizoram | Shifting cultivators | Assam-Burmese, Kuki-Chin |
| Mara | 90 | Mizoram | Shifting cultivators | Assam-Burmese, Kuki-Chin |
| Lai | 92 | Mizoram | Shifting cultivators | Assam-Burmese, Kuki-Chin |
| Lusei | 92 | Mizoram | Shifting cultivators | Assam-Burmese, Kuki-Chin |
| Nepali | 220 | Sikkim | Agriculture | Tibeto-Himalayan, Himalayan |
| Bhutia | 150 | Sikkim | Farming | Tibeto-Himalayan, Bhotia |
| Lepcha | 96 | Sikkim | Farming | Tibeto-Himalayan, Himalayan |
| Naga | 152 | Manipur | Shifting cultivators | Assam-Burmese, Naga |
| Kuki | 150 | Manipur | Shifting cultivators | Assam-Burmese, Kuki-Chin |
| Muslims* | 160 | Manipur | Trade | Assam-Burmese, Kuki-Chin |
| Garo | 160 | West Bengal | Shifting cultivators | Assam-Burmese, Bodo |
| Chinese | 222 | China | Farming | Sinitic subfamily |
| Malayasian Javanese | 218 | Malaysia | Trade | Austronesian family, Malayo-Polynesian, Western Malayo-Polynesian, Sundic, Javanese |

*Western Eurasian origin.
**All populations belong to Sino-Tibetan family (except Malaysian Javanese) and Tibeto-Burman subfamily (except Chinese).

eastern wing of the Himalayan mountain ranges around Pasighat, Arunachal Pradesh. High molecular weight DNA was isolated using the standard phenol/chloroform method (Sambrook et al. 1989). The extracted DNA was then quantitated using the Quantiblot® Human DNA Quantification Kit (Applied Biosystems, Foster City, CA, USA).

Individual Adi Pasi DNA samples were amplified for the 15 microsatellite loci: D5S818, FGA, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, D2S1338, D19S433, vWa, TPOX, D18S51 using AmpF*I* STR® Identifiler™ Multiplex system (Applied Biosystems). The amplified products were then separated on a 4% polyacrylamide gel using the ABI Prism 377 automated DNA sequencer (Applied Biosystems). The resultant data was analyzed using Gene scan analysis software (Version 3.7) and the allele designation was done with Genotyper DNA fragment analysis (Version 3.7) software (Applied Biosystems).

The allele frequencies of 12 STR loci viz., D5S818, FGA, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, vWa, TPOX, D18S51 for the remaining 13 populations (other than Adi Pasi of Arunachal Pradesh) and the locus-wise heterozygosity values of 12 populations (except Chinese) were obtained by a literature survey (Kashyap et al. 2002; Law et al. 2002; Trivedi et al. 2002; Kashyap et al. 2003; Maity et al. 2003; Kashyap et al. 2004b).

*Statistical analysis*

Based on 12 loci, the locus-wise genetic diversity ($G_{ST}$) (Nei 1973, 1987) and pair-wise genetic distances between populations using the modified Cavalli-Sforza distance ($D_A$) and the standard genetic distance ($D_{ST}$) measures of Nei et al. (1983) were computed using the software DISPAN (Ota 1993). We have also estimated $F_{ST}$ distances between populations to investigate if a similar genetic relationship is obtained. Subsequently two phylogenetic trees: the unweighted pair group method with arithmetic mean (UPGMA) tree and neighbor-joining (NJ) tree were constructed based on $D_A$, $D_{ST}$ and $F_{ST}$ distance measures. Since $D_A$ distance measure is the most efficient for obtaining correct phylogenetic trees under various evolutionary conditions and also is least affected by small sample size (Takezaki and Nei 1996), our discussions are based on this distance measure. For confirmation of the robustness of the clustering pattern obtained for the studied TB populations, we have included four additional Dravidian populations, three (Tanjore Kallar, Irular and Iyengar Brahmin) from southern India and one (Dheria Gond) from central India and constructed an NJ-phylogenetic tree based $D_A$ distances. To check the consistency of the clustering pattern of the dendrograms, a total of 1000 and also 10 000 bootstrap replications were separately performed. Principal Component Analysis (PCA) was also performed for the TB populations based on the $F_{ST}$ distance matrix using SPSS software, Chicago, IL, USA. The PCA plot may be more robust than the dendrogram clustering method, especially when bootstrap values are low and the similar clustering in both the PCA plot and the dendrogram indicates the consistency of the results obtained. The regression model of R-matrix analysis (Harpending and Ward 1982) was applied to the STR data for the 13 populations excluding the Muslim population of Manipur because the model assumes a common origin or stock from which all other subpopulations have branched out in the course of time, primarily, as a result of the relative influence of gene flow and genetic drift. It is assumed that the Tibeto-Burman populations are possibly derived from a common stock and are supposed to have come from the southern part of China or Tibet in the remote past. Muslims of Manipur are of non-Asian origin and hence do not share ancestry with the other studied populations. MM is a distinct (morphologically and ethnically) endogamous group separate from the TB populations and, as such, they are expected to stand outside the cluster of TB populations both in the dendrogram and the R-matrix plot. Therefore, R-matrix analysis was also performed for all the

14 populations including MM to validate their separate identity and their clear separation in dendrogram.

## Results

### Genetic diversity among the studied populations

The heterozygosity values (locus-wise and population wise) including the overall average heterozygosity and the locus-wise $G_{ST}$ values for the 13 TB populations are given in Table II. The average heterozygosity that reflects within-population heterogeneity exhibits a narrow range between 0.68 (among the Lepcha of Sikkim) and 0.79 (among the Hmar of Mizoram). The locus-wise $G_{ST}$ values range from 2.2% at THO1 to 8.4% at D18S51. The average $G_{ST}$ value is 4.7%, which suggests a high degree of differentiation among the Tibeto-Burman speaking populations at the studied loci.

### Genetic affinities

The Nei's $D_A$ distance matrix is presented in Table III. The geographically distant populations, especially Lai of Mizoram and Muslims of Manipur show maximum distance whereas the two local regional populations Hmar and Lusei of Mizoram show the minimum distance value. The Chinese population shows lower distance value with the Indian TB Populations than the Javanese.

Although we have constructed both the UPGMA and the NJ trees based on Nei's $D_A$, $D_{ST}$ and $F_{ST}$ distance matrices, we present here the NJ dendrogram based on Nei's $D_A$ distance (Figure 2), as both $D_A$ and $F_{ST}$ dendrograms depict the same pattern of relationships among the studied populations. The tree based on $D_{ST}$ distance measure shows a slightly different pattern of clustering and since it is not a robust distance measure it was not considered. The $D_A$-NJ tree consists of three major subclusters. Cluster I consists of Hmar, Mara, Lai and Lusei of Mizoram and Adi Pasi of Arunachal Pradesh. Cluster II includes the Nepali, Bhutia and Lepcha of Sikkim. Chinese and Malaysian Javanese and Kuki of Manipur are seen as outliers to these two clusters. Naga of Manipur and Garo of West Bengal form a separate subcluster. Muslims of Manipur stand out of the tree suggesting that they have least genetic affinity with the studied populations. The highest boot strap value obtained was 63 for Malaysian Javanese, the remaining bootstrap values being low.

Figure 3 shows the clustering pattern of the 14 studied (TB) populations along with four Dravidian speaking communities. There are four subclusters. All the Tibeto-Burman populations show a similar clustering pattern to that obtained in Figure 2 except the Chinese and Javanese samples, which form a separate subcluster along with the Mizoram and Arunachal samples. Similarly Kuki clusters along with Naga and Garo. The three Dravidian population samples from southern region form a separate cluster, whereas the Dheria Gond, a central region Dravidian tribe stands outside of the cluster. Similar to Figure 2, the MM, a non-TB population, separate out as outlier from the rest of the samples. The highest boot strap value (62) is observed in the case of the two Dravidian populations from the southern region, the other boot strap values being low.

Figure 4 shows the PCA plot of 14 populations (including MM group) based on Component I and Component II scores. In general, the scattering pattern obtained shows similarities with the clustering pattern obtained in the dendrogram (Figure 2). Of all

Table II. Locus-wise and population-wise heterozygosity and locus-wise $G_{ST}$ values for the studied populations.

| Locus | Adi Pasi | Hmar | Mara | Lai | Lusei | Nepali | Bhutia | Lepcha | Naga | Kuki | Garo | Muslims | Javanese | $G_{ST}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D18S51 | 0.680 | 0.850 | 0.822 | 0.826 | 0.848 | 0.843 | 0.844 | 0.812 | 0.869 | 0.895 | 0.674 | 0.712 | 0.844 | 0.084 |
| D21S11 | 0.803 | 0.825 | 0.867 | 0.783 | 0.783 | 0.921 | 0.652 | 0.819 | 0.885 | 0.775 | 0.888 | 0.811 | 0.844 | 0.042 |
| D5S818 | 0.765 | 0.825 | 0.778 | 0.783 | 0.783 | 0.713 | 0.717 | 0.773 | 0.792 | 0.834 | 0.783 | 0.889 | 0.826 | 0.030 |
| D13S317 | 0.799 | 0.900 | 0.822 | 0.826 | 0.761 | 0.778 | 0.688 | 0.750 | 0.700 | 0.821 | 0.797 | 0.701 | 0.780 | 0.043 |
| D7S820 | 0.620 | 0.875 | 0.844 | 0.826 | 0.783 | 0.855 | 0.719 | 0.637 | 0.667 | 0.709 | 0.696 | 0.778 | 0.725 | 0.049 |
| vWA | 0.815 | 0.800 | 0.778 | 0.608 | 0.804 | 0.810 | 0.750 | 0.796 | 0.667 | 0.834 | 0.798 | 0.889 | 0.826 | 0.043 |
| D8S1179 | 0.827 | 0.850 | 0.822 | 0.761 | 0.674 | 0.858 | 0.813 | 0.682 | 0.792 | 0.792 | 0.674 | 0.667 | 0.670 | 0.042 |
| D3S1358 | 0.696 | 0.700 | 0.733 | 0.652 | 0.609 | 0.667 | 0.750 | 0.728 | 0.584 | 0.874 | 0.631 | 0.500 | 0.670 | 0.049 |
| THO1 | 0.600 | 0.425 | 0.556 | 0.565 | 0.543 | 0.715 | 0.657 | 0.432 | 0.715 | 0.773 | 0.736 | 0.806 | 0.734 | 0.022 |
| CSF1PO | 0.670 | 0.825 | 0.689 | 0.630 | 0.652 | 0.742 | 0.750 | 0.660 | 0.765 | 0.781 | 0.782 | 0.630 | 0.706 | 0.049 |
| TPOX | 0.555 | 0.700 | 0.674 | 0.674 | 0.674 | 0.634 | 0.657 | 0.546 | 0.669 | 0.524 | 0.591 | 0.763 | 0.606 | 0.056 |
| FGA | 0.795 | 0.900 | 0.800 | 0.913 | 0.848 | 0.800 | 0.563 | 0.523 | 0.792 | 0.542 | 0.826 | 0.834 | 0.881 | 0.063 |
| Average | 0.719 | 0.790 | 0.765 | 0.737 | 0.730 | 0.778 | 0.713 | 0.680 | 0.741 | 0.763 | 0.740 | 0.748 | 0.772 | 0.0477 |

Table III.   Matrix of Nei's $D_A$ distance between pairs of populations studied.

| Population | Adi Pasi | Hmar | Mara | Lai | Lusei | Nepali | Bhutia | Lepcha | Naga | Kuki | Garo | Muslims | Chinese | Javanese |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Adi Pasi | – | | | | | | | | | | | | | |
| Hmar | 0.0582 | – | | | | | | | | | | | | |
| Mara | 0.0907 | 0.0714 | – | | | | | | | | | | | |
| Lai | 0.0596 | 0.0448 | 0.0923 | – | | | | | | | | | | |
| Lusei | 0.0639 | 0.0405 | 0.0808 | 0.0533 | – | | | | | | | | | |
| Nepali | 0.0734 | 0.067 | 0.1009 | 0.0637 | 0.063 | – | | | | | | | | |
| Bhutia | 0.1306 | 0.1098 | 0.1456 | 0.1164 | 0.1167 | 0.1066 | – | | | | | | | |
| Lepcha | 0.1109 | 0.1115 | 0.1227 | 0.0958 | 0.113 | 0.0951 | 0.1474 | – | | | | | | |
| Naga | 0.0863 | 0.11 | 0.1301 | 0.1001 | 0.114 | 0.1192 | 0.1599 | 0.1467 | – | | | | | |
| Kuki | 0.1053 | 0.0965 | 0.1105 | 0.1047 | 0.1025 | 0.1178 | 0.1566 | 0.142 | 0.0815 | – | | | | |
| Garo | 0.1418 | 0.1518 | 0.1747 | 0.1536 | 0.1562 | 0.1394 | 0.1898 | 0.1792 | 0.0782 | 0.1108 | – | | | |
| Muslims | 0.2665 | 0.276 | 0.2921 | 0.2946 | 0.2531 | 0.2766 | 0.2945 | 0.2883 | 0.2675 | 0.2649 | 0.2437 | – | | |
| Chinese | 0.0518 | 0.0444 | 0.0725 | 0.0444 | 0.0427 | 0.0528 | 0.097 | 0.0946 | 0.0668 | 0.0757 | 0.1131 | 0.2462 | – | |
| Javanese | 0.0568 | 0.0536 | 0.0771 | 0.0552 | 0.0526 | 0.0637 | 0.1001 | 0.0981 | 0.0824 | 0.0765 | 0.1248 | 0.2419 | 0.0297 | – |
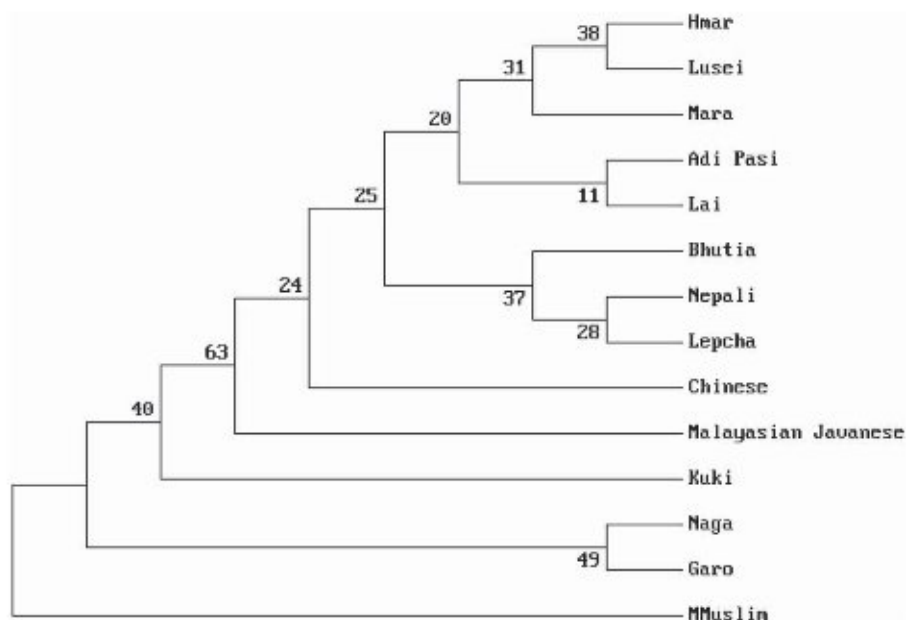
Figure 2. Neighbor-joining tree constructed on $D_A$ distance matrix depicting the genetic relationship between the 14 populations, based on 12 microsatellite markers.
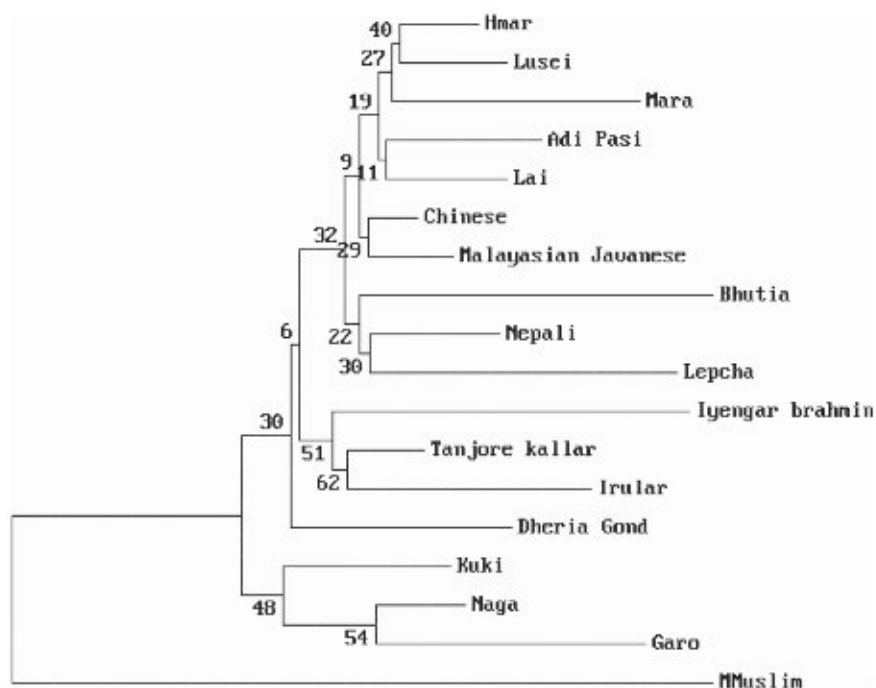


Figure 3. Neighbor-joining tree constructed on $D_A$ distance matrix depicting the genetic relationship between the 14 Tibeto-Burman populations and four Dravidian populations, based on 12 microsatellite markers.
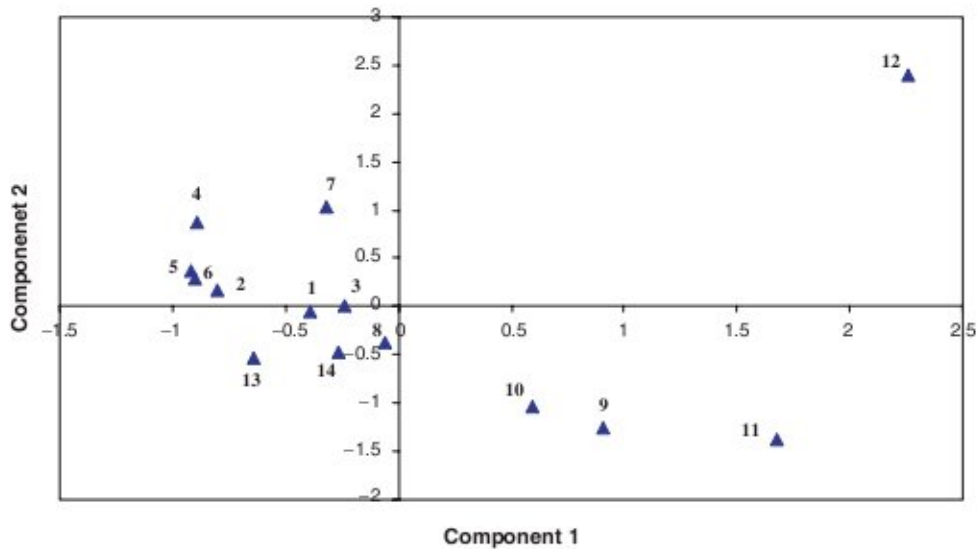
Figure 4. PCA plot of 14 populations on first two components based on $F_{ST}$ distance matrix. Codes used in the figure: 1- Adi Pasi, 2- Hmar, 3- Mara, 4- Lai, 5- Lusei, 6- Nepali, 7- Bhutia, 8- Lepcha, 9- Naga, 10- Kuki, 11- Garo, 12- Manipuri Muslims 13- Chinese, 14 -Malaysian Javanese.

the populations the MM group is placed in the first quarter separated distinctly from other TB populations confirming the dendrogram obtained. The second quarter consists of five TB populations; Bhutia from Sikkim is placed away whereas Nepali is closely placed along with Hmar, Lai and Lusei (Mizoram populations) and the third quarter represents diverse groups: Chinese, Malaysian Javanese along with Lepcha from Sikkim, Adi Pasi from Arunachal Pradesh and Mara from Mizoram. The fourth quarter shows clustering of three populations; Naga and Kuki of Manipur and Garo of West Bengal, whereas in the dendrogram Naga and Garo form a close cluster and Kuki clusters with rest of the TB group.

*Patterns of gene flow*

The two evolutionary forces, genetic drift and gene flow, play crucial roles in the micro evolutionary process among subdivided populations which differentiate from a common stock. Apart from morphological features, folklore and other cultural materials (such as beads) indicate that the studied populations may have belonged to some common stock in the remote past and that the genetic differentiation among the diverse regional populations could be attributed to the relative influence of genetic drift and gene flow. To investigate the influence of these two evolutionary forces, we have used the Harpending and Ward regression model. According to the model, if the subdivided populations experience an equal extent of external gene flow and random genetic drift, there will be a linear trend between average heterozygosity and distance from the centroid. The populations receiving greater than the average gene flow from outside will be more heterozygous and less influenced by drift and therefore lie above the theoretical regression line. Conversely, populations experiencing less than average gene flow from outside show lower
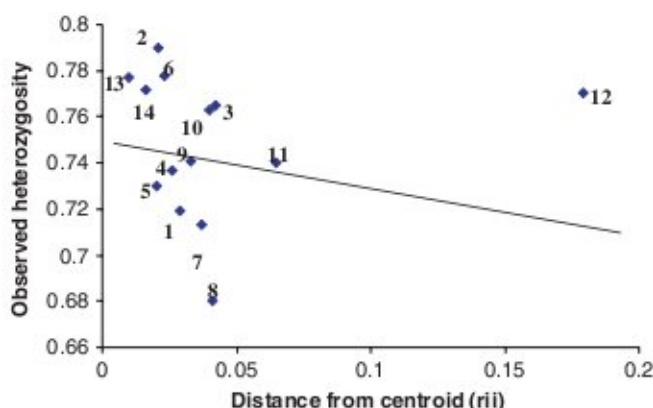
Figure 5. Regression plot of the average heterozygosity *vs.* the distance from centroid of the studied populations. Codes used in the figure: 1- Adi Pasi, 2- Hmar, 3- Mara, 4- Lai, 5- Lusei, 6- Nepali, 7- Bhutia, 8- Lepcha, 9- Naga, 10- Kuki, 11- Garo, 12- Manipuri Muslims 13- Chinese, 14- Malaysian Javanese.

heterozygosity values and are influenced by genetic drift to a greater extent and hence lie below the regression line. In the regression plot (Figure 5) the populations of Lepcha, Bhutia, Adi Pasi, Lai, Lusei and Naga are positioned below the theoretical regression line and Hmar, Nepali, Kuki, Mara lie above the regression line along with the two global populations, the Chinese and Javanese. The Garo also lie above the regression line but are separated from the other populations and the MM group, being an outlier, is placed distantly from the rest of the 13 TB populations.

## Discussion

Tibeto-Burman populations differ from other major linguistic families of India in several respects. They form a separate ethnic group but constitute several diverse groups and subgroups. They are supposed to have arrived at different time periods and from different directions and settled mostly in the high Himalayan mountainous terrain that stretches from north to northeastern parts of the country. Some populations have remained isolated and retained their identity which is reflected in their unique genetic profiles. However, others have adopted local customs and mixed with neighboring local castes and tribes thereby exhibiting vast socio-cultural and genetic diversity as evident from some classical genetic markers (Roychoudhury 1981, 1992; Kumar et al. 2004). In this regard, studies of these populations based on classical genetic markers reveal some interesting results, e.g., complete lack of $A_2$, cde, K, $p^c$, and $AK^2$ genes, lack of isozyme ALDH-1 (Roychoudhury and Nei 1997), a high prevalence (about 50%) of lactase malabsorption (Flatz 1987), low frequency of AIBG*2 allele (Juneja et al. 1989), high frequency of G6PD deficiency in Naga (Seth and Seth 1971), absence of 'Gd⁻' variant in Adi and Hmar and high frequency of this variant in Bodos (Saha et al. 1990). Continuing from classical genetic observations, this study also reveals a few unique alleles and a few rare alleles (Krithika et al. 2005). The study also shows low average heterozygosity for the STR loci in the studied populations when compared to other regional populations, castes and some tribes. This can be attributed to high endogamy and little gene flow among the studied populations. The high average $G_{ST}$

value of 4.7% among the studied populations reflects the population structure characteristics of small population size and high endogamy and isolation that in turn may have resulted in rapid genetic differentiation.

One of the objectives of our study is to utilize the microsatellite data to investigate the biological validity and verification of the possible diverse stock of splinter groups that have branched out from a common origin in the remote past. The phylogenetic tree obtained from the microsatellite data supports the historical data of their possible origin and migration. Historical accounts suggest that the Kuki and Naga tribes of Manipur have migrated from Burma. The separation of Kuki and Naga from others in the phylogenetic tree is an indication of their separate migration history. The tribes of Arunachal and Mizoram have migrated from the southern region of Tibet (Roychoudhury 1992; Roy 1997; Nath 2000) and accordingly they form a distinct subcluster. The Sikkim tribes had a separate migration history and represent a different subcluster. Thus the clustering pattern of the 14 regional populations obtained reflects their diversities and geographical contiguity. An interesting feature is the clustering of the Garo of West Bengal with the far away Manipur populations instead of clustering with the geographically proximate Sikkim populations. Similar clustering of Garo with the Manipur tribes in particular has been noted in classical genetic markers and STR markers as well (Dutta et al. 2002). Garo in West Bengal live along with caste populations of Indo-European speakers and are exposed to external gene flow. Similarly, Naga and Kuki tribes of Manipur (geographically close to Burma and Assam) have experienced admixture with the local populations in the regions adjoining their settlement. This could be a plausible reason for the separate clustering of these populations. This clustering pattern gets further support when four Dravidian populations are included. Where the clustering pattern of TB populations remains unchanged, the Dravidian populations form a separate cluster reflecting their geographic proximity. The low bootstrap values observed in the dendrograms can be attributed to the heterogeneity of the populations and the nature of the loci studied. Although these values give less statistical significance to the clustering pattern obtained, the consistent pattern obtained by PCA plot and by different phylogenetic methods suggest the underlying genetic affinity of these populations. This is also the case in other such studies and in other populations as well. Hence bootstrap values are not the only benchmark criterion to depend upon but consistency of results, obtained by other methods (e.g., PCA plot), is an indication of the biological validity of the studied populations.

In general, genetic affinity and diversity among regional populations show certain patterns with respect to geographic contiguity, ethnicity, language, sociocultural affiliation, and past history (Roychoudhury 1981; Roychoudhury 1992; Cavalli-Sforza et al. 1994; Majumder 1998a,b). Based on classical genetic markers among eastern Tibeto-Burman populations, Roychoudhury (1992) found geographical proximity to be a major determinant factor of genetic affinity between eastern and northeastern populations irrespective of their linguistic or ethnic affiliations in the past. In our present study we have verified that this observation, based on biochemical markers, is supported by the microsatellite data as well! Our results show that overall, geographic contiguity is a major influencing factor of genetic affinity among diverse populations with varied migrational backgrounds and distributed over a wide geographical area. Geographically proximate regional populations with common recent history or from the same region or from a common stock cluster together suggesting close genetic affinity. Since microsatellite markers are supposed to reflect population structure and microevolutionary trend irrespective of the geographic affinity of the populations, the geographical clustering of regional populations reflects their recent past genetic history, supporting their antiquity and the historical accounts of their

possible migration from Burma, in the case of Manipur populations, and from Tibet, in the case of Arunachal Pradesh, Mizoram and Sikkim populations. The Garo of West Bengal live alongside other caste populations and as a result they may have experienced external gene flow from the local non-Tibeto-Burman speakers of non-Asiatic origin, as can be inferred from their position above the regression line of the Harpending and Ward model. This could be one of the reasons for their possible deviant position observed in the phylogenetic tree.

Microevolutionary process among subdivided populations is primarily guided by genetic drift and gene flow. The relative roles of these two forces vary depending on the structure of the populations. Some Tibeto-Burman populations are isolated whereas some others live along with local caste populations in urban areas. Isolated populations should reflect least gene flow and show low levels of heterozygosity and other populations should show high levels of heterozygosity due to possible admixture with local populations that promotes greater external gene flow. The results of the regression plot show that populations placed below the regression line (Adi Pasi, Lepcha, Bhutia, Lai, Lusei and Naga) reflect low levels of heterozygosity as a result of least external gene flow. This is attributed to their remote location and high endogamy, whereas populations that lie above the regression line (Hmar, Nepali, Kuki, Mara, Chinese, Malaysian Javanese and Garo) show high heterozygosity values due to greater than average external gene flow possibly resulting from their admixture with neighboring populations. Manipur Muslim population being a non-TB population is genetically a separate ethnic group and this is reflected in both the dendrogram and the regression plot where it is an outlier, distinctly separated from the rest.

There has been a considerable debate as to whether Tibeto-Burman populations of Asian origin came from the Tibetan region from the north or from Burma through the southeastern route. Earlier workers have tried to understand the possible migration route of these populations based on their language, cultural similarity, and folklore (Roychoudhury 1992; Blackburn 2004). The historical records do support the geographic contiguity hypothesis invoking common origin and antiquity; the populations from Manipur and Tripura especially show their antiquity to Burma and populations in the northern fringe to Tibet. Due to 'limited number of common possible loci' available for the populations investigated,' Roychoudhury (1992), however, was cautious about inferring the biological validity of their affinity and their possible antiquity. In this regard, the dendrogram and PCA plot obtained based on microsatellite loci indeed supports their diverse routes of historical migration and affinity with the populations across the northern and eastern borders.

## Acknowledgements

# References

Basu A, Mukherjee N, Roy S, Sengupta S, Banerjee S, Chakraborty M, Dey B, Roy M, Roy B, Bhattacharyya NP, Roychoudhury S, Majumder PP. 2003. Ethnic India: A genomic view, with special reference to peopling and structure. Genome Res 13:2277–2290.

Bhasin MK, Walter H, Danker-Hopfe H. 1992. The distribution of genetical, morphological and behavioral traits among the peoples of Indian origin. Delhi: Kamla-Raj Enterprises.

Bhattacharyya N, Basu P, Das M, Pramanik S, Banerjee R, Roy B, Roychoudhury S, Majumder PP. 1999. Negligible gene-flow across ethnic boundaries in India, revealed by analysis of Y-chromosomal polymorphisms. Genome Res 9:711–719.

Blackburn S. 2004. Memories of migration: Notes on legends and beads in Arunachal Pradesh, India. European Bulletin of Himalayan Research 25/26:15–60.

Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. Nature 325:31–36.

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton, NJ: Princeton University Press.

Chakrabarti CS, Roy M, Sengupta NK, Lalthantluanga R, Majumder PP. 2002. Genetic relationships among some tribal groups inhabiting the north eastern, eastern and sub Himalayan regions of India. Ann Hum Genet 66:361–368.

Cordaux R, Saha N, Bentley GR, Aunger R, Sirajuddin SM, Stoneking M. 2003. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. Eur J Hum Genet 11:253–264.

Das BM, Deka R, Das R. 1980. Haemoglobin in six populations of Assam. J Indian Anthrop Soc 15:153–156.

Deka R, Reddy AP, Mukherjee BN, Das BM, Banerjee S, Roy M, Dey B, Malhotra KC, Walter H. 1988. Haemoglobin E distribution in ten endogamous population groups of Assam, India. Hum Hered 38:261–266.

Dani AH. 1960. Prehistory and protohistory of eastern India. Calcutta: Firma KL Mukhopadhyay.

Dutta R, Reddy BM, Chattopadhyay P, Kashyap VK, Sun G, Deka R. 2002. Patterns of genetic diversity at the nine forensically approved STR loci in the Indian populations. Hum Biol 74:33–49.

Elwin V. 1959. A Philosophy for NEFA. Shillong: North-East Frontier Agency.

Flatz G. 1987. Genetics of lactose digestion in humans. Adv Hum Genet 16:1–77.

Gadgil M, Joshi NV, Shambu Prasad UV, Manoharan S, Suresh patil 1903. 1997. The Indian human heritage. Hyderabad: University Press.

Grierson GA. 1909. Linguistic survey of India. Vol. III Tibeto-Burman Family, Part I. Calcutta: Superintendent Government Printing.

Harpending HC, Ward R. 1982. Chemical systematics and human evolution. In: Nitecki M, editor. Biochemical aspects of evolutionary biology. Chicago IL: University of Chicago Press. pp 213–256.

Juneja RK, Saha N, Gahne B, Tay JS. 1989. Distribution of plasma alpha-1-B-glycoprotein phenotypes in several Mongoloid populations of east Asia. Hum Hered 39:218–22.

Kashyap VK, Chattopadhyay P, Dutta R, Vasulu TS. 2004a. Genetic structure and affinity among eight ethnic populations of eastern India: Based on 22 polymorphic DNA loci. Am J Hum Biol 16:311–327.

Kashyap VK, Sarkar N, Sahoo S, Sarkar BN, Trivedi R. 2003. Genetic variation at fifteen microsatellite loci in human populations of India. Curr Sci 85:464–473.

Kashyap VK, Guha S, Trivedi R. 2002. Concordance study on 15 STR loci in three major populations of Himalayan state Sikkim. J Forensic Sci 47:1163–1167.

Kashyap VK, Ashma R, Gaikwad S, Sarkar BN, Trivedi R. 2004b. Deciphering diversity in populations of various linguistic and ethnic affiliations of different geographical regions of India: Analysis based on 15 microsatellite markers. J Genet 83:49–63.

Kivisild T, Tolk HV, Parik J, Wang Y, Papiha SS, Bandelt HJ, Villems R. 2002. The emerging limbs and twigs of the east Asian mt DNA tree. Mol Bio Evol 19:1737–1751.

Krithika S, Trivedi R, Kashyap VK, Vasulu TS. 2005. Validation of 15 microsatellite loci for detection and human identification in Adi Pasi, a small subtribe of Adi tribal cluster of Arunachal Pradesh, India. Leg Med 7:306–310.

Kumar V, Basu D, Reddy BM. 2004. Genetic heterogeneity in northeastern India: Reflection of tribe-caste continuum in the genetic structure. Am J Hum Biol 16:334–345.

Law MY, To KY, Ho SH, Pang BCM, Wong LM, Wun SK, Chan KL. 2002. STR data for the Powerplex™ 16 loci for the Chinese population in Hong Kong. Forensic Sci Int 129:64–67.

Maity B, Nunga SC, Kashyap VK. 2003. Genetic polymorphism revealed by 13 tetrameric and 2 pentameric STR loci in four Mongoloid tribal populations. Forensic Sci Int 132:216–222.

Majumdar DN. 1980. Northeast India: A profile. In: Sharma TC, Majumdar DN, editors. Eastern Himalayas: A study on anthropology and tribalism. N. Delhi: Cosmo.

Majumder PP. 1998a. People of India: Biological diversity and affinities. Evol Anthrop 6:100–110.

Majumder PP. 1998b. Genes, diversities, and peoples of India. In: Macer DRJ, editor. Ethical challenges as we approach the end of the human genome project. Eubios Ethics Institute. pp 20–33.

Malhotra KC, Vasulu TS. 1993. Structure of Human populations in India. In: Majumder PP, editor. Human population genetics. New York: Plenum. pp 207–233.

Nath J. 2000. Cultural heritage of tribal societies. Vol. 1 (The Adis). 1st ed. New Delhi: Omsons Publications.

Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 19:153–170.

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

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Ota T. 1993. Dispan: Genetic distance and phylogenetic analysis. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park: Pennsylvania.

Papiha SS, Chahal SMS, Mastana SS. 1996. Varialibility of genetic markers in Himachal Pradesh, India: Variation among subpopulations. Hum Biol 68:629–654.

Rapson EJ. 1955. People and Languages. In: Rapson EJ, editor. Cambridge history of India. Vol. 1. Ancient India, Delhi: S. Chand, pp 33–57.

Roychoudhury AK, Nei M. 1997. The Emergence and dispersal of Mongoloids. J Indian Anthrop Soc 32:01–469.

Roychoudhury AK. 1981. The genetic composition of the people in Eastern India. J Indian Anthrop Soc 16:153–170.

Roychoudhury AK. 1992. Genetic relationships of the populations in eastern India. Ann Hum Biol 19:489–501.

Roy S. 1997. Aspects of Padam Minyong culture. 3rd ed. Guwahati: M/S Purbadesh Mudran.

Ruhlen M. 1991. A guide to world's languages Stanford, CA: Stanford University Press.

Saha N, Hong SH, Wong HA, Tay JSH. 1990. Red cell glucose-6-phosphate dehydrogenase phenotypes in several Mongoloid populations of eastern India: Existence of a non deficient fast variant in two Australasian tribes. Ann Hum Biol 17:529–532.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: A laboratory manual cold spring harbor. New York: Cold Spring Harbor Laboratory Publications.

Seth PK, Seth S. 1971. Biogenetical studies of Nagas: G-6-P-D deficiency in Angami Nagas. Hum Biol 43:557–561.

Singh KS. 1998. People of India: India's communities. Vol. 1. New Delhi: Oxford University Press.

Takezaki N, Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite data. Genetics 144:389–399.

Trivedi R, Chattopadhyay P, Maity B, Kashyap VK. 2002. Genetic polymorphism at nine microsatellite loci in four high altitude Himlayan desert human populations. Forensic Sci Int 127:50–55.

Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RR, Oefiner PJ, Cavalli-sforza. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet 65:43–62.

Walter H, Mukherjee BN, Gilbert K, Lindenberg P, Dannewitz A, Malhotra KC, Das BM, Deka R. 1986. Investigations on the variability of haptoglobin, transferring and Gc polymorphisms in Assam, India. Hum Hered 36:388–396.

WHO 1964. Research in population genetics of primitive groups. WHO Tech Rep Sr. p 279.

**Résumé.** *Arrière plan:* Les populations tibéto-birmanes (TB) sont l'un des quatre groupes linguistiques majeurs de l'Inde. On considère qu'elle appartiennent à des stocks différents et présentent de fortes variations en culture et en langue, bien que leur distance génétique, leur ancienneté et l'histoire de leur migration aient été peu étudiées. Des études génétique moléculaires permettront peut être de clarifier leur antiquité et leur diversité génétique.

*But:* Cet article cherche à comprendre l'étendue et la magnitude des affinités génétiques et des diversités de 14 populations TB (12 indiennes et deux groupes globaux), à éprouver les résultats fondés sur les marqueurs génétiques classiques et vérifier les récits historiques de leur migration et leur histoire génétique fondée sur 12 marqueurs microsatellites.

*Sujets et méthodes:* Les fréquences alléliques de 12 loci STR appartenant à 13 populations asiatiques (tibéto-birmanes), ont été extraites de la littérature et les données Adi Pasi ont été obtenues par le typage microsatellitaire de leurs échantillons sanguins. Les 12 loci étudiés sont: D5S818, FGA, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, vWa, TPOX, D18S51. la relation génétique de ces populations a été analysée au moyen de trios calculs de distance différents, deux arbres phylogénétiques et trois représentations factorielles.

*Résultats:* L'hétérozygotie moyenne varie de 68 à 79% et la valeur moyenne de GST est 4,7%. Le dendrogramme établi à partir des distances DA montrent le regroupement des populations sur la base de leur diversité et de leur contiguïté géographique; Les populations du Mizoram et de l'Arunachal Pradesh se regroupent, les populations du Sikkim forment un sous groupe individualisé et les populations Manipur ainsi que les Garo de l'ouest du Bengale se séparent des autres regroupements. Le modèle de régression d'Harpending et Ward montre les populations isolées placées en dessous de la ligne de régression, tandis que celles qui ont reçu un flux génique externe sont placées au-dessus. Les résultats sont en accord avec les récits migratoires du folklore concernant leur antiquité possible avec les populations du Tibet et de Chine méridionale.

*Conclusion:* La contiguïté géographique en général, ponctuée de barrières d'isolement est un facteur majeur influençant les affinités génétiques des populations TB; les populations contiguës dans une région montrent de plus grandes relations génétiques qu'elles n'en ont avec les populations TB d'un plus vaste ensemble géographique. Le résultat de l'étude des microsatellites confirme également l'histoire des diverses routes de migration de ces populations.


**Zusammenfassung.** *Hintergrund:* Die Tibetisch-Burmanischen (TB) Völker sind eine der vier großen linguistischen Bevölkerungsgruppen Indiens. Man nimmt an, dass sie unterschiedlicher Herkunft sind und eine große kulturelle und sprachliche Vielfalt aufweisen; allerdings sind ihre genetische Beziehung, ihre Frühzeit und die Geschichte ihrer Wanderungen unter den regionalen Völkern wenig untersucht. Es wird erwartet, dass molekulargenetische Studien die Frühzeit und die genetische Diversität dieser Völker aufklären.

*Ziel:* Diese Arbeit versucht, das Ausmaß und die Größenordnung der genetischen Verwandtschaft und Diversität unter 14 TB Völkern (12 Indischen und 2 Gruppen aus der übrigen Welt) zu verstehen, ferner die Ergebnisse auf der Grundlage klassischer genetischer Marker zu untersuchen und anhand von 12 Mikrosatelliten-Markern die historischen Berichte ihrer Wanderungen und ihre genetische Geschichte zu verifizieren.

*Probanden und Methoden:* Daten zur Allelhäufigkeit für 12 STR loci von 13 Asiatischen (Tibetisch-Burmanischen) Populationen wurden der Literatur entnommen, und die Adi Pasi Daten wurden über Mikrosatelliten-Typisierung von Blutproben gewonnen. Die 12 untersuchten loci sind D5S818, FGA, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, vWa, TPOX, D18S51.

Drei verschiedene Entfernungsmesser, zwei phylogenetische Stammbäume und PCA-Plots wurden verwendet, um die genetische Beziehung der untersuchten Völker zu verstehen.

*Ergebnisse:* Die Werte für mittlere Heterozygosität schwanken von 68 bis 79% und der mittlere $G_{ST}$-Wert ist 4,7%. Das auf $D_A$-Abständen basierende Dendrogramm zeigt eine Ungleichverteilung (clustering) von Völkern je nach Diversität und geographischer Nähe; besonders die Mizoram und Arunachal Pradesch Völker erweisen sich als besonders nah verwandt, Völker aus Sikkim bilden eine separate Untergruppe, und die Völker aus Manipur zusammen mit den Garo aus West Bengalen sind am weitesten von den anderen getrennt. Das Harpending und Ward Regressionsmodell zeigt isolierte Populationen unterhalb der Regressionslinie, und andere, die von außen hereingekommenem genetischen Einfluss ausgesetzt waren, oberhalb dieser Linie. Die Ergebnisse passen zu Volkssagen über gemeinsame frühzeitliche Völkerwanderungen mit Tibetischen und Südchinesischen Populationen.

*Zusammenfassung:* Insgesamt ist die geographische Nähe, betont durch isolierte Barrieren, ein wesentlicher Einflussfaktor für genetische Verwandtschaft unter den TB Völkern; benachbarte Völker innerhalb einer Region zeigen eine höhere genetische Verwandtschaft als geographisch weit voneinander entfernt lebende Populationen. Das Ergebnis der Mikrosatellitenstudie bestätigt auch die verschiedenen historischen Völkerwanderungsrouten.


**Resumen.** *Antecedentes:* las poblaciones Tibetano-Burman (TB) son uno de los cuatro principales grupos lingüísticos de población de la India. Se considera que pertenecen a diferentes linajes y muestran una gran variación en lengua y cultura; sin embargo, sus relaciones genéticas, su antigüedad y la historia migratoria entre poblaciones regionales, han sido poco investigadas. Se espera que los estudios genéticos moleculares muestren claramente la antigüedad y la diversidad genética de estas poblaciones.

*Objetivo:* este artículo intenta comprender la extensión y magnitud de las afinidades y diversidades genéticas entre 14 poblaciones TB (12 indias y 2 grupos globales), investigar los hallazgos basados en marcadores genéticos clásicos y verificar los relatos históricos de su migración e historia genética, basándose en 12 marcadores de microsatélites.

*Sujetos y métodos:* se obtuvieron datos sobre las frecuencias alélicas para 12 loci STR de 13 poblaciones asiáticas (Tibetano-Burman) a partir de la literatura; los datos de los Adi Pasi se obtuvieron mediante el análisis de microsatélites a partir de sus muestras sanguíneas. Los 12 loci estudiados fueron D5S818, FGA, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, vWa, TPOX, D18S51. Para comprender la relación genética entre las poblaciones estudiadas, se utilizaron tres medidas de distancia diferentes, dos árboles filogenéticos y la representación de un ACP.

*Resultados:* los valores medios de heterozigosis oscilan entre 68 y 79% y el valor GST medio ha sido de 4,7%. El dendrograma, basado en la distancia DA, muestra la agrupación de las poblaciones según su diversidad y contigüidad geográfica. Las poblaciones de Mizoram y Arunachal Pradesh se agrupan especialmente juntas, las poblaciones de Sikkim forman un subgrupo separado y las poblaciones de Manipur y las Garo de Bengala occidental se separan de los otros grupos. El modelo de regresión de Harpending y Ward muestra que las poblaciones aisladas se sitúan por debajo de la línea de regresión y el resto, que han experimentado un flujo génico externo, se sitúan por encima de la línea. Los resultados obtenidos apoyan los relatos tradicionales sobre antiguas migraciones de las poblaciones tibetanas y del sur de China.

*Conclusiones:* en conjunto, la contigüidad geográfica, interrumpida por barreras de aislamiento, es un importante factor que influye en la afinidad genética entre las poblaciones TB. Las poblaciones próximas dentro de una región muestran una mayor relación genética entre ellas que con las poblaciones TB distantes, dentro de un área geográfica extensa. Los resultados del estudio de los microsatélites también apoyan la historia de las diferentes rutas de migración de estas poblaciones.