Geographic Contiguity, Patterns of Gene Flow and Genetic Affinity among the Tribes of Arunachal Pradesh, India

S. Krithika, S. Maji and T.S. Vasulu*

Biological Anthropology Unit, Indian Statistical Institute, Kolkata 700108, West Bengal, India

ABSTRACT The study examines the regional genetic diversity among 23 Arunachal Pradesh tribes based on 2 loci (ABO and PTC). The results show wide variation in allele frequencies. The 'r' allele shows higher frequency (than 'p' and 'q') and show geographical variation. The results of NJ tree and PCA plot show separation of tribal groups that fairly corresponds to their geographical locations and ethno-historical backgrounds. The Harpending and Jenkins regression plot suggests that these tribes are getting differentiated primarily due to genetic drift and genetic isolation, where gene flow plays a significant role in a few tribes. Also, the affinity among the regional groups based on their ethno-historical origin and migration and genetic diversity was considered by a model-based approach especially by Rao's hierarchical analysis. The results of the study thus support ethno-historical accounts of their antiquity and possible common origin.

INTRODUCTION

Understanding the pattern of biological variation and its association with geography, language, ethnicity and culture of Indian populations has been the focus of study in anthropological genetics. Empirical studies indicate variable patterns of association among castes and tribes inhabiting wide geographic regions (Bamshad et al. 1996; Basu et al. 2003; Bhasin 1992; Chakrabarti et al. 2002; Cordaux et al. 2003). For example, Dravidian and Austro-Asiatic speaking tribes inhabiting different geographic regions show wide genetic diversity thus supporting the hypothesis of their heterogeneous origin, geographic isolation and migration history (Basu et al. 2003; Cordaux et al. 2003; Kumar et al. 2003). Whereas geographically proximate tribes and sub tribes within a region reflect close genetic affinity irrespective of their cultural and linguistic differences (Deka et al. 1988, 1999; Gaikwad et al. 2006). In this regard such patterns of association between languages, culture and geography is not fully explored, except for the overall regional study among the Tibeto-Burman speaking Mongoloid ethnic populations of northeast (Das 1968; Das et al. 1980; Dutta et al. 2002; Krithika et al. 2006; Kumar

Corresponding Author: T.S.Vasulu, Biological Anthropology Unit, Indian Statistical Institute, 203 BT Road, Kolkata 700108, West Bengal, India.
Telephone: +91-33-25753215
E-mail: vasulu@isical.ac.in, vasulu@gmail.com

et al. 2004). For example, a study based on classical genetic markers across regional populations showed genetic similarity with respect to geographical proximity, irrespective of their origin and linguistic affiliation (Roychoudhury 1981, 1992). Whether similar such patterns can be observed within regional tribes and sub tribes are yet to be explored.

In this regard it is interesting to investigate the geographic contiguity hypothesis within region or a particular state in northeast. This is because, in northeast, each state is unique in terms of geographical location, diversity of populations that differ in culture, language, subsistence strategy, origin, history and migration pattern etc (Dani 1960; Elwin 1959; Nath 2000). However there is hardly any region specific biological study of a particular state (Adak 2004). These regional studies are of significance since it can give better clarity of their confounding hypotheses concerning their origin, antiquity and migration (based on folklore, cultural materials and historical anecdotes). Our study concentrates on the genetic variation in Arunachal Pradesh in northeast region. Of all regions in northeast, Arunachal Pradesh presents an illustrative example of wide geographic area, diverse populations of a variety of tribes with possible common origin, antiquity from Tibet but differential migration and dispersal routes and settlement.

To understand the extent of biological affinity and diversity among the regional tribes of this

state, we explored the different available published data sets on classical genetic markers (Bhasin et al. 1992, Singh et al. 1994). Our study includes 23 tribes from different districts of Arunachal Pradesh and is based on 2 classical genetic markers, namely, ABO blood groups and Phenyl Thio Carbamide (PTC) tasting ability. We have investigated the regional variation, affinity and diversity among different tribes, the relative roles of genetic drift and gene flow that is responsible for their differentiation and the influence of ethno-historical migration of these regional tribes based on a model bound approach.

MATERIALS AND METHODS

Data: The available allele frequency data of nine classical genetic markers for different tribes of Arunachal Pradesh were obtained by manual data mining through literature survey (Bhasin et al. 1992; Singh et al. 1994). However, due to some data constraints (the number of populations versus the number of common loci), we finally could obtain and investigate the genetic affinity of 23 tribes and sub tribes studied for at least two common loci (ABO & PTC). The details of the populations studied and their geographic distributions are given in (Table 1). These populations were arranged according to their geographic locations from northwestern to southeastern region as suggested by Elwin (1958), first anthropologist to study the regional diversity of Arunachal Pradesh tribes based on geography, religion, origin and language.

While exhausting the available published data on Arunachal Pradesh populations based on classical genetic markers, certain difficulties and limitations were encountered regarding the availability of the suitable data type, data reliability, consistency and accuracy. We excluded those data in which the allele frequencies did not add up to 1 and also those populations whose sample size was below 50.

Affinity and Diversity Analysis: Based on 2 loci (ABO & PTC), the locus wise genetic diversity (GST) (Nei 1973, 1987) and pair-wise genetic distances between 23 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). Subsequently two phylogenetic trees: the unweighted pair group method with arithmetic mean (UPGMA) tree and neighbor-joining (NJ) tree were constructed based on the two distance measures, using software Mega 2.0 (Kumar et al. 2001). 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 only on this distance measure. To investigate the consistency of the dendrogram obtained, Principal Component Analysis (PCA) was done based on the D_A distances between populations using the SPSS 11.0 software, Chicago, IL. Further, we have performed the Harpending and Jenkins R-matrix analysis (Harpending and Jenkins 1973) for the studied populations to investigate the relative roles of genetic drift and gene flow. To assess the relative amount of gene flow into each population, the regression model (Harpending and Ward 1982) was applied. Also, the expected heterozygosity values of each population were computed (Relethford and Blangero 1990) following which a regression plot was drawn between observed and expected heterozygosity values of populations to understand the genetic differentiation between the studied populations.

The hypothesis of ethno-historical migration was investigated by Rao's hierarchical analysis of apportionment of diversity (Rao 1982; Majumder 1988). It basically involves in estimating the contribution of within and between diversities to the total diversity of the populations at different hierarchical levels defined on the basis of ethno-history and other information related to population structure. The association between the geographic distribution and the allele frequencies of the studied loci of these populations was calculated by Spearman's rank correlation.

RESULTS

Genetic Diversity: The tribes show wide variation in ABO frequencies thereby exhibiting regional variation; O being more frequent than A or B in majority of them. The Central and Eastern tribes show preponderance of O blood group over A or B whereas the Northern Tibetan Buddhists show higher frequency of A than O. The lowest frequency of A is found among Monpa Kalaktang and the highest (64%) among Mishimi Digaru. Tangsa and Khampti in Eastern and Central region show minimum frequency of

A and Khowa from Kameng shows the maximum frequency for A and minimum for B. In case of AB blood group, Aka shows lowest and Monpa Kalaktang shows the highest frequency. For PTC taste sensitivity most of the tribes show higher percentage of tasters (84 to 94%) except among Northern Tibetan Buddhists (70% to 77%). The average heterozygosity values range from 39.6 in Aka to 55.0 or 56.0 among the Northern Monpa tribes.

All the populations show higher frequency of r allele (O) and except two (Khampti and Tangsa) the rest of the populations show higher p (A) than q (B) allele frequencies. The frequency of r allele varies from 0.495 (Monpa Kalaktang) to 0.70 (Mishimi Digaru), Mishimi Digaru and Tangsa show lowest frequency of p allele (0.14 and 0.13 respectively) and Apatani Guchi and Khowa show the highest value (0.33). The q allele is observed to be lowest among Khowa (0.034) and highest among Monpa Tawang (0.244), both located in Kameng river valley. In general, PTC tasters allele (T) frequency is higher than nontasters (t) in all the populations except in two populations of Monpa (Kalaktang and Dirang) that also show a low nontaster frequency (45% and 46%). The highest taster allele frequency was observed among Aka (79%) and Khowa (78%). The gene diversity (G_{ST}) per loci among the 23 tribes show a high value for ABO and least for PTC.

Genetic Affinities: The D_A distance matrix of

Table 1: Geographical locations, linguistic background, religion and subsistence pattern of the populations studied

Populations	Geographical Location (District)	Language/ Dialect	Religion	Subsistence pattern
Monpa-Dirang	Kameng	Monpa	Buddhism (Mahayana)	Agriculture
Monpa-Kalak Tang	Kameng	Monpa	Buddhism (Mahayana)	Agriculture
Monpa-Tawang	Kameng	Monpa	Buddhism (Mahayana)	Agriculture
Sherdukpen	Kameng	Sherdukpen	Buddhism (Mahayana)	Agriculture
Aka	Kameng	Hrusso	Donyie-Polo	Agriculture- Shifting Cultivation
Nishi	Kameng	Nishi	Donyie-Polo	Agriculture- Shifting Cultivation
Khowa	Kameng	Bagun	Buddhism	Agriculture- Shifting Cultivation
Miji	Kameng	Sajalong	Donyie- Polo	Agriculture- Shifting Cultivation
Tagin	Upper Subansiri	Tagin	Donyie- Polo	Hunting and Gathering
Hillmiri	Upper Subansiri	Miri	Donyie-Polo	Agriculture- Shifting Cultivation
Apatani-Guchi	Lower Subansiri	Apatani	Donyie-Polo	Agriculture- Shifting Cultivation
Apatani-Guth	Lower Subansiri	Apatani	Donyie-Polo	Agriculture- Shifting Cultivation
Adi	Siang	Adi	Donyie-Polo	Hunting and Gathering, Shifting Cultivation
Galong	Siang	Adi	Donyie-Polo	Hunting and Gathering, Shifting Cultivation
Adi- Minyong	Siang	Adi	Donyie-Polo	Hunting and Gathering, Shifting Cultivation
Mishimi-Digaru	Lohit	Mishimi	Donyie- Polo	Agriculture- Shifting Cultivation
Mishimi-Idu	Lohit	Mishimi	Donyie-Polo	Agriculture- Shifting Cultivation
Mishimi-Miju	Lohit	Mishimi	Donyie-Polo	Agriculture- Shifting Cultivation
Khampti	Lohit	Khamprtishan	Buddhism (Hinayana)	Agriculture, Trade
Singpho	Lohit	Singpho	Buddhism (Hinayana)	Agriculture
Tangsa	Tirap	Tangsa	Animists	Hunting and Gathering
Wancho	Tirap	Wancho	Raung-Baurang	Hunting and Gathering
Naga-Nocte	Tirap	Nocte	Kat Jauban- Wang Jauban, Vaishnanavism	Agriculture

23 populations based on 5 alleles (2 loci) show maximum distance value between Monpa Dirang and Khowa (0.0501), followed by the populations Monpa Tawang and Khowa (0.0457) whereas Apatani-Guchi and Sherdukpen show the minimum distance value (0.0001) followed by Mishmi Idu and Galong (0.0004).

For investigating the pattern of clustering showing the affinity among the 23 populations, phylogenetic trees were constructed. Although we have constructed both the UPGMA and the NJ trees based on Nei's D_A and D_{ST} distance matrices, we present here only the NJ dendrogram based on Nei's D_A distance (Fig.1), since D_{ST} distance measure is not a robust distance measure and both UPGMA and NJ dendrograms (based on D_A) depict similar pattern of relationships among the studied populations. The D_A-NJ tree shows two major clusters and each consists of two subclusters. The Northern Kameng group of Aka, Khowa, Miji, Sherdukpen along with Nishi, Adi Minyong and Mishimi Idu

form one major cluster and Northen Tibetan Buddhist group of Monpa tribes and Burmese Buddhist group and two Mishimi group form another major cluster. The first cluster has two sub clusters: first sub cluster comprising of populations Khowa, Aka, Miji, Apatani-Guchi, Sherdukpen and Apatani-Guth; Wancho standing out of this sub cluster and the second sub cluster consisting Nishi, Galong, Adi-Minyong and Mishimi Idu populations. Tagin population is observed to be an outlier to cluster 1. The second major cluster also includes two sub clusters. The first sub cluster consist of 3 sub tribes of Monpa tribe (Dirang, Kalaktang, Tawang), Singpho, Adi and the second sub cluster comprises of the populations Tangsa, Khampti, Naga Nocte, Mishimi Miju and Mishimi Digaru. Hill Miri is seen as an outlier to this second major cluster.

The principal component analysis (PCA) plot (Fig.2) based on D_A distance matrix shows clustering pattern similar to the pattern obtained from the NJ tree except for Hill Miri which gets
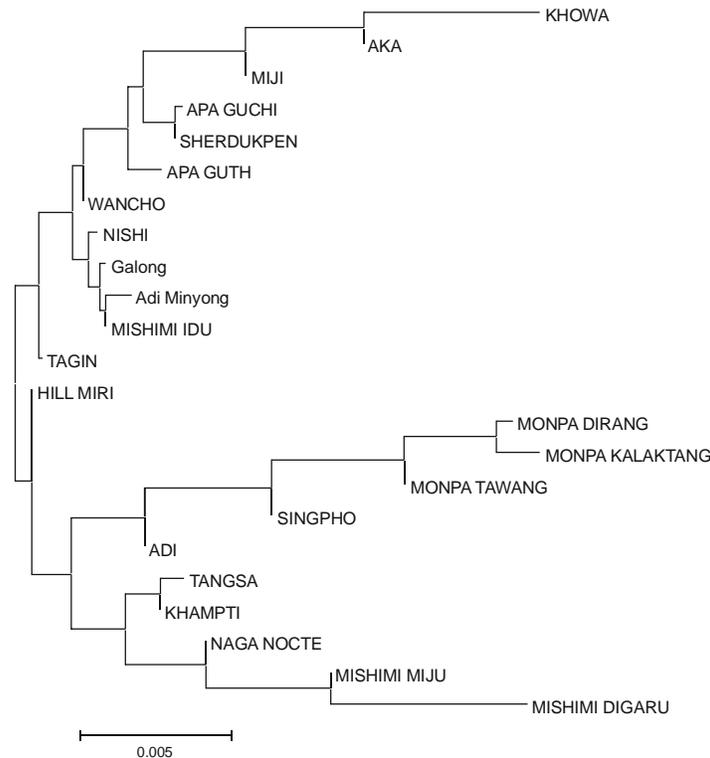

Fig. 1. Neighbor-joining tree constructed on D_A distance matrix depicting the genetic relationship between the 23 populations, based on 2 classical genetic markers.
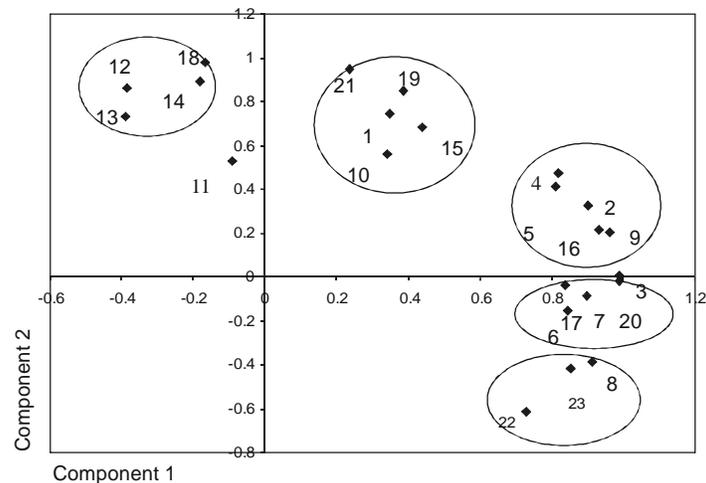
Fig. 2. PCA plot of 23 populations on the first two components derived from the principle component analysis based on the D_A distance matrix. 1, Adi; 2, Adi Minyong; 3, Galong; 4, HillMiri; 5, Tagin; 6, Apatani-Guchi; 7, Apatani-Guth; 8, Miji; 9, Mishimi-Idu; 10, Mishimi-Miju; 11, Mishimi-Digaru; 12, Monpa-Dirang; 13, Monpa-Kalaktang; 14, Monpa-Tawang; 15, Naga-Nocte; 16, Nishi; 17, Sherdukpen; 18, Singpho; 19, Tangsa; 20, Wancho; 21, Khampti; 22, Khowa; 23, Aka.

clustered with Tagin and other populations of cluster I instead of clustering with populations of cluster II. Although Mishimi Digaru and Mishimi Miju lie within the same cluster, they are distantly placed in the plot unlike as in NJ tree where they form a tight cluster.

Patterns of Gene Flow: Genetic drift and gene flow are the two crucial forces that govern the microevolutionary process among sub divided populations getting differentiated from a common stock. To understand the relative effect of these two evolutionary forces, we have used the Harpending and Ward regression model. According to this model, if the sub divided populations experience equal extent of external gene flow and random genetic drift, there will be a linear trend between average heterozygosity and distance from the centroid under the assumption of constant effective population size and same migration rate between populations. The populations receiving greater than the average gene flow from outside will be more heterozygous and least influenced by drift and therefore lie above the theoretical regression line. Conversely, populations experiencing lesser than average gene flow from outside show lower heterozygosity values and influenced by genetic

drift to a greater extent and hence lie below the regression line. In the regression plot (Fig. 3), a majority of the studied populations are positioned at the vicinity of theoretical regression line except, Aka, Khowa and Mishimi Digaru that lie below the regression line but are separated distinctly from other populations near the regression line. Whereas Monpa Tawang, Monpa Dirang, Monpa Kalaktang and Adi populations lie above the regression line distantly situated from other populations in the plot. Further distribution of 23 populations with respect to observed and expected heterozygosity values were investigated for confirmation of the above results. Except a few populations (Mishmi-Miju (10) and Mishmi-Digaru (11), a majority of the populations are placed very close to the expected line (Figure not included).

Ethno-history and Hierarchical Analysis: The hypothesis of ethno-historical migration and possible origin suggesting the relationship among the Arunachal tribes can be tested by hierarchal analysis of genetic diversity. The available historical information on their oral tradition, folklore and cultural materials, suggest that these tribes basically came from either Tibet-China or from Myanmar (Burma) and occupied different geographic locations at different time

Fig. 3. Regression plot of the observed heterozygosity versus the distance from centroid of the populations studied. 1, Adi; 2, Adi Minyong; 3, Galong; 4, Hill Miri; 5, Tagin; 6, Apatani-Guchi; 7, Apatani-Guth; 8, Miji; 9, Mishimi-Idu; 10, Mishimi-Miju; 11, Mishimi-Digaru; 12, Monpa-Dirang; 13, Monpa-Kalaktang; 14, Monpa-Tawang; 15, Naga-Nocte; 16, Nishi; 17, Sherdukpen; 18, Singpho; 19, Tangsa; 20, Wancho; 21, Khampti; 22, Khowa; 23, Aka.

periods, though we do not have the clarity when and from where they have arrived. It is known that Northwest tribes have come from adjoining Tibet as early as 17th century whereas the tribes in eastern region came from Burma during the 18th century. The tribes in central regions, especially the Tani (linguistic) groups trace their origin from Tibet. Their geographical location and their possible relationship based on their ethno-historical migration are depicted in a clustering tree (Fig.4). Their depicts different levels of relationship based on past migration and possible origin of the regional tribes. Five different levels of clustering showing the affinity between the regional studied tribes can be recognized. These are: H_0 for the possible common origin or stock in the remote past. H_1 *(TBC)* denoting their place of migration, H_2 representing the North-Central group *(NCG)*, H_3, the regional groups *(RG)* and H_4 designating the local populations *(LP)* under each RG. The TBC has two main branches and they have been identified as Tibeto-Chinese Cluster (H_{TCC}) and Eastern Burmese Cluster (H_{EBC}). The H_{TCC} has two main sub clusters, *viz.*, Northwest Kamang Group (H_{NKG}) and Central Tani Group (H_{CTG}). Both these trace their origin from Tibet and occupy different

regions in the Kamang river valley and in Central river valleys. Further, the NKG has two regional clusters that identify them according to their geographic location and religious practices and language: Tibeto-Buddhist Group (TBG) and Animist Group (ANG) and the Central Tani group (CTG) has four clusters: Tagin-Hillmiri Group (THG), Apatani Group (APG), Adi Group (ADG) and Mishimi Group (MSG). The H_{EBC} consists of two sub clusters: Burmese Buddhist Group (BBG) and Naga Tribal Group (NTG). The different levels of clustering and its contribution to the total genetic diversity of the region can be considered by hierarchical gene diversity analysis. The apportionment of diversity at various levels for each loci and overall contribution is shown in Table 2. The diversity between the studied populations when compared to the total, based on two loci, is on an average 4.7% and it is higher in case of PTC (7.3%) than ABO loci. Of all the levels of comparisons of different clusters of Arunachal Pradesh tribes, the results show a maximum diversity at the regional level (RG) when compared to the total diversity. About 45% diversity is accounted at the regional level between the regional clusters of tribes and this is 69% and 37.5% in case of
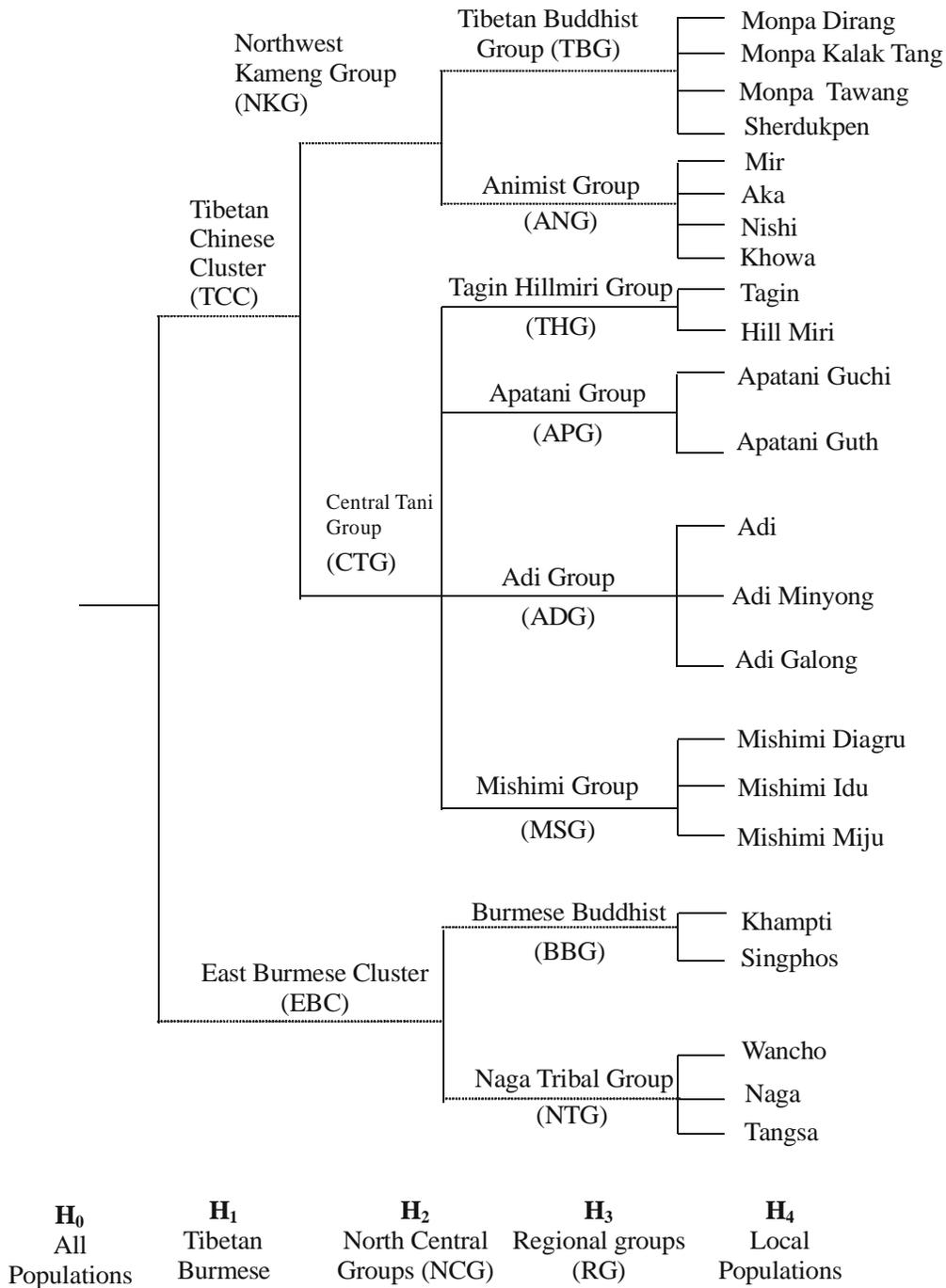
Fig. 4. Schematic diagram showing hierarchical levels of clustering of 23 Arunachal Pradesh tribes based on geographic location, place of origin, religion and language.

Table 2: Apportionment of diversity among 23 studied population groups based on 2 loci

Due to	Formula	ABO	%	PTC	%	Overall	%
Between All Population - Tibetan Burmese Clusters(TBC)	H0-H1	0.00114	52.87	0.00059	20.00	0.00087	34.11
Tibetan Burmese Clusters(TBC) - North Central Group (NCG)	H1-H2	0.00101		0.00235		0.00168	
All Population - North Central Group (NCG)	H0-H2	0.00216	28.13	0.00295	23.79	0.00255	25.39
NCG – Regional Groups (RG)	H2-H3	0.00552		0.00945		0.00748	
All Population – RG	H0-H3	0.00768	65.92	0.0124	37.58	0.01004	44.98
RG – Local Population (LP)	H3-H4	0.00397		0.02059		0.01228	
All Population – LP	H0-H4	0.01165	2.15	0.03299	7.27	0.02232	4.69
Local Population (LP)	H4	0.53047		0.42025		0.47536	
Total	H0	0.54213		0.45325		0.49769	

ABO and PTC respectively. The northern central group (NCG) in Kameng river valley show about 25% diversity whereas between the combined two major clusters (TBC) consisting of northern (NCG) and eastern groups (Patkai hills) (EBC) of tribes show greater diversity (34%) and it is mostly due to ABO loci (53%) than PTC. A greater regional diversity is expected in view of the geographical isolation and differentiation among the sub tribes. This might also suggest geographical patterns with respect to genetic traits. The correlation analysis suggests significant positive association between the geographical location and genetic diversity in case of r allele ($r_s = 0.62$) and negative in case of p allele ($r_s = -0.71$) of ABO blood group and is not significant in case of PTC.

DISCUSSION

Northeastern parts of India are conglomeration of several Tibeto-Burman speaking tribes of diverse origin; however, the biological diversity of these populations is far from complete. Though there were studies investigating the biological diversity of the entire northeastern region together, there is a dearth of regional studies in different states or regions. In genetic point of view, such regional studies will help to investigate some of the intriguing problems about their origin and affinity with other populations. Arunachal Pradesh is inhabited by several major tribes and sub tribes geographically distributed in northern, central and eastern parts of the region. It is known that they have come from neighboring Tibetan and (or) Burmese regions in the past, however, their original place and time and route of migration and genetic affinity and diversity with other tribes is not clear. In general, fission and fusion process is characteristic population structure of early tribal groups. This is expected to result (besides other factors) in wide inter-regional biological diversity between the major tribes. The results of the study indicate wide diversity in allele frequencies and average heterozygosity values of the studied loci. This is one of the expectations of the tribes differentiated by the process of fission-fusion dynamics. Besides it could be attributed to other population structure variables especially, of small population size, geographic isolation and least external gene flow, which promotes local genetic differentiation. This was supported by the Harpending and Jenkins regression model. The results of the regression plot do support the least influence of gene flow especially among Khowa (22), Aka (23) in East Kameng river valley in northern region and among Mishimi-Diagaru (11) and in at least another six tribes in central and eastern regions where they all lie below the line suggesting greater influence of genetic drift than gene flow.

What is interesting is whether the available data is fairly enough to investigate the regional diversity and support their antiquity of possible common origin? The results obtained do indicate that this was sufficient enough to infer the regional affinity among the different regional populations. The clustering pattern as seen from the NJ tree and PC plot fairly separates the regional populations with some exception. The northern groups including Tibetan Buddhist and eastern groups including Burmese Buddhists cluster and the central Tani group form separate clusters suggesting their possible common antiquity and common origin in the remote past.

Even the results obtained from model bound approach of hierarchical analysis based on ethno-historical information also show higher levels of diversity (i.e. about 50% total diversity) between the regional groups, whereas the diversity between the local sub tribes is very low. Possibly additional information on several tribes is expected to provide more clarity of the regional variation. In general, overall recent developmental activities are fast eroding their traditional tribal life style, which might promote the process of fusion in major tribes that may eventually result in fast changes of their genetic make-up. At a later date, such studies will become difficult or perhaps not possible because of some recent trends of increase in trade and changes in life style activities. In this regard these populations provide opportunities for studies in molecular genetics to investigate the debatable issues of their origin, migration history and clues to health and disease aspects.

ACKNOWLEDGEMENT

REFERENCES

Adak DK 2004. Variation in anthropometric characters and their relationship with climatic factors: A study on the peoples of Arunachal Pradesh. *Anthropol Anz (Germany)*, **62:** 445-61.

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

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 Region.* Delhi: Kamla-Raj Enterprises.

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.

Dani AH 1960. *Prehistory and Protohistory of Eastern India.* Calcutta: Firma KL Mukhopadhyay.

Das BM 1968. ABO blood groups in the tribal populations of northeast India with special reference to the Khasi. *Anthropologists*, **Special Vol:** 85-92.

Das BM, Deka R, Das R 1980. Haemoglogin 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.

Deka R, Shriver MD, Yu LM, Heidrecich EM, Jin L Zhong Y, McGarvey ST, Agarwal SS 1999. Genetic variation at twenty three microsatellite loci in sixteen human populations. *J Genet*, **78:** 99-121.

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.

Gaikwad S, Vasulu TS, Kashyap VK 2006. Microsatellite diversity reveals the interplay of language and geography in shaping genetic differentiation of diverse proto-australoid populations of West-Central India. *Am J Phy Anthrop,* **129:** 260-267.

Harpending HC, Jenkins T 1973. Genetic distance among South African populations. In: MH Crawford and PL Workman (Ed.): *Methods and Theories of Anthropological Genetics.* Albuquerque: University of Mexico press pp. 177-199.

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

Krithika S, Trivedi R, Kashyap VK, Vasulu TS 2006. Antiquity, geographic contiguity and genetic affinity among Tibeto-Burman populations of India: a microsatellite study. *Ann Hum Biol,* **33(1):** 26-42.

Kumar S, Tamura K, Jakobsen IB, Nei M 2001. MEGA 2: molecular evolutionary genetics analysis software, bioinformatics. *Bioinformatics,* **17:** 1244-1245.

Kumar V, Reddy BM 2003 Status of Austro-Asiatic groups in the peopling of India: An exploratory study based on the available prehistoric, linguistic and biological evidences. *J Biosci,* **28:** 507-522.

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.

Majumder PP 1988. Analysis of genetic diversity in subdivided population. In: Malhotra KC (Ed.): *Statistical Methods in Human Population Genetics.* Calcutta: Eka press pp. 276-310.

Nath J 2000. *Cultural Heritage of Tribal Societies, Vol I (The Adis).* 1st ed. New Delhi: Omsons Publications.

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

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

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.

Rao CR 1982. Diversity: Its measurement, decomposition, apportionment and analysis. *Sankhya,* **44:** 1-22.

Relethford JH, Blangero J 1990. Detection of differential gene flow from patterns of quantitative variation. *Hum Biol,* **62:** 5-26.

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.

Singh KS, Bhalla V, Kaul V 1994. *Biological Variation in Indian Populations.* Calcutta: Oxford University Press.

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