

Journal: <u>www.joqq.info</u> Originally Published: Volume 4, Number 1 (Spring 2008) Reference Number: 41.003

GENETIC STRUCTURE OF AN ISOLATED SUB-TRIBE OF THE ADI PEOPLE OF ARUNACHAL PRADESH STATE IN NORTHEAST INDIA: ISONYMY ANALYSIS AND SELECTIVE NEUTRALITY OF A SURNAME DISTRIBUTION IN ADI PANGGI

Author(s): Suvendu Maji and T.S. Vasulu

Genetic Structure of an Isolated Sub-Tribe of the Adi People of Arunachal Pradesh State in Northeast India: Isonymy Analysis and Selective Neutrality of Surname Distribution in Adi Panggi

Suvendu Maji and T. S. Vasulu

Abstract

The distribution of surnames was studied to infer the population structure aspects of migration and genetic drift and the expectation of surname neutrality of progeny size distributing among the Panggi (Pangi), a small isolated Tibeto-Burman sub-tribe of the Adi tribe, subsisting on hunting-and-gathering in Upper Siang district of Arunachal Pradesh, India. Random isonymy (I), Karlin-McGregor's v and Fisher's α were estimated. Log-log and Pareto distributions were also studied for the occurrence of surname distribution among husbands, wives for testing of surname neutrality. The estimates of homonymy vary between villages and show greater variation among husbands than wives. The log-log distributions of surname occurrence in the patrilocal tribe show a nonlinear trend and the pattern differs in husband and wife samples. The results also show a non-linear trend of surname distribution with respect to progeny size among the post- reproductive and reproductive wives and their husbands. The surname analysis among the Adi Panggi provide some insights about the genetic structure of the isolated tribe which will help the molecular genetic studies to unravel the differential paternal and maternal past genetic history of the tribe.

Introduction

Surnames are a unique bio-cultural trait, which provide a convenient means of investigating microevolution in human populations. The patrilineal mode of inheritance of surnames mimics highly polymorphic genes on the Y-chromosome; the non-biological nature of its dispersal is expected to be independent of fertility and mortality differentials and therefore satisfies the expectations of the neutral theory of evolution (Kimura, 1980), which has been described by Karlin and MacGregor (1967) as the theoretical distribution of different mutant forms that are maintained in a population under the equilibrium between random genetic drift, mutation, and migration. The occurrence of different surnames in human populations also conforms to Fisher's logarithmic (Chakraborty et al, 1981) and Pareto's discrete (Fox 1983) distributions.

The earliest application of theoretical models to the surname distribution has been considered among the parishes in the Parma valley (Yasuda et al, 1974) and in island populations of Sardinia (Zei et al, 1983a, b). Since then, there has been progress in the use of isonymy studies to investigate the genetic structure among a wide variety of populations, particularly in Europe and Latin American countries (Barrai et al, 1987; Barrai et al, 2003; Colantonio et al, 2003). These studies have shown that the logarithmic distribution of surname frequency follows a linear trend (Barrai et al, 1987) in conformity with the neutral allele model (Kimura, 1983); however, a few studies have shown deviation from the expected linear trend (Barrai et al, 1987; Barrai et al, 2002; Barrai et al, 2003). These isonymy studies have revealed (a) geographical patterns in surname distributions as a result of differential migration as demonstrated in Ferrara, an immigrant community, Italy (Barrai et al, 1987; Barrai et al, 1989; Barrai et al, 1990), Austria (Barrai et al, 2000), Perugia (Rodriguez et al, 1993), Spain (Rodriguez et al, 2003); (b) the reflection of social and natural selection in Denmark (Boldsen 1992); (c) deviation from linearity as a result of excesses of surname repetition especially, in Netherlands (Barrai et al, 2002), Sicily (Scapoli et al, 1997) and Belgium (Barrai et al, 2003). The majority of the cited studies are based on large sample sizes consisting of an entire nation or region; however, such studies among small isolated populations have rarely been conducted. In this regard, it will be of interest to examine the expectation of neutrality of surname distribution in small populations, especially among isolated tribes, since the demographic events and marriage practices bring rapid changes along the kinship lines influencing the surname structure and in such situation the neutrality of surname distribution may not be expected.

This study describes the surname distribution and examines the neutral allele model in an isolated small tribe *viz.*, Adi Panggi--one of the several sub-tribes of the Adi tribal cluster--inhabiting the Upper Siang River Valley in central Arunachal Pradesh State, India.

Address for correspondence: T. S. Vasulu, vasulu@gmail.com

Received: 26 Oct 2007; accepted: 05 Feb 2008.

Materials and Methods

Adi Tribe

The Adi Tribe consists of several sub-tribes inhabiting different altitudes of the southeastern part of Himalayan mountain terrain along the Siang river valley in the central region of the State of Arunachal Pradesh (Roy 1960; Singh 1994; Blackburn 2004; Lego 2005). In general, an Adi tribal village consists of a group of families belonging to a few specific clans living together at different locations of the valley. Adi Panggi is one of the smallest isolated sub-tribes. They reside in 7 villages situated in different valleys or hill slopes over an area of about 50 square km in Geku Circle, Upper Siang district, and number about four thousand individuals (Koley 2005). The sub-tribe speaks the Adi language, a member of the Tibeto-Burman linguistic family. The northern part of the region shares a border with China. Although the Adi sub-tribes share a common historical migration, possible common origin, and linguistic and cultural affiliations, each sub-tribe forms a separate group that maintains its identity. Each sub-tribe is geographically isolated, practices high endogamy, and has specific clans and surname structure different from other sub-tribes of the Adi (Roy 1960; Lego 2005; Koley 2005). Like other tribes in India, they are patrilocal, patrilineal, and patriarchal: sons tend to stay in the village or nearby villages, while daughters migrate to their husband's village after their marriage and adopt their husband's surname (Das 1953).

The clan and surname is indicative of their putative origin from their possible common ancestral stock and indicates genetic kinship. The surname structure plays an important role in their marriage, warfare and in hunting and cultural activities and are stable over generations. These factors make surname analysis useful for investigating the genetic structure of the population. There are only a few studies on the Adi Panggi, e.g., ABO polymorphism (Bhattacharjee 1954; Krithika et al, 2006), cultural aspects (Sharma 1960), anthropometric variation (Roy 1966) and a recent ethnographic study (Koley 2005) have been reported.

Sample

Demographic data and blood samples of Adi Panggi (Pangi) tribal population were collected from six villages in Geku circle (Figure 1) for a molecular population genetics study among the Adi tribe of Arunachal Pradesh, India. The study was approved by the ethical committee of Indian Statistical Institute, Kolkata. The genetic analysis will be the subject of an upcoming article.

For the surname analysis, the surname of the husband and the 'maiden surname' of the wife were collected through pedigree data from a field survey in 2006 (Maji et al. 2007). The surname data were collected from 154 husbands and 130 wives. Of the seven villages, six villages were studied for surname distribution: Sumsing



Figure 1. Map showing the location of different Panggi villages (•) in the Upper Siang District along the Siang River Valley in Arunachal Pradesh, India

(SS) and Sibum (SB) are remotely located around 15-30 km away from Geku Town (GT) whereas the remaining three villages, Ramku (RK), Kumku (KK) and Peram (PR), are located close to GT.

Isonymy Analysis

Random isonymy (I) was calculated from the frequency distribution of the abundance of each surname separately in males (p_i) and females (q_i) and for both in each (ith) village. If p_i and q_i are the frequencies of a particular surname in males (husbands) and females (wives) respectively, then the random isonymy estimate among males, females and both males and females is:

$$I = p_i^2$$
 or $I = q_i^2$ and $I = p_i q_i$

Since the random isonymy is biased for the samples' size variations, an unbiased isonymy and its variance was estimated (Dyke et al, 1983; Relethford 1988; Barrai et al, 1989; Barrai et al, 1991). If there are N_i males and N_i females in the samples the unbiased random isonomy I_{ij} is defined as:

$$\mathbf{I}_{ij} = \left(\sum n_{is} \bullet n_{js}\right) / \mathbf{N}_i \bullet \mathbf{N}_j,$$

where, n_{is} and n_{js} are the numbers of individuals with surname *s* in populations i and j, respectively, and N_i and N_j are the total number of surnames in populations i and j, and summation is over all s surnames.

Surname Distribution and Neutral Allele Model

Since the surnames mimic a genetic trait on the Y-chromosome, and the surname distribution in a large population also correlates with genetic diversity (Barrai et al, 1996), it is expected to conform, under certain assumptions, to the selection neutrality of an infinite allele model described by Karlin and MacGregor (1967), which can be accounted for by the logarithmic distribution first given by Fisher. That is, if we let S represent the number of times that a surname occurs in a population, and let K represent the frequency of S (i.e., K is the number of surnames occurring this same number of S times), then we may characterize the distribution by graphing log K versus log S. Karlin and MacGregor's v and Fisher's α were estimated for the Adi Panggi after the formula proposed by Zei et al. (1983a, 1983b) and Barrai et al. (1992) respectively, where

$$\alpha = 1/I_{ij}$$
 and $\mathbf{v} = \alpha/(N_i + \alpha)$

and I is random isonymy, N_i is the total number of individuals and α is a measure of the surname diversity and **v** is a measure of the migration into the population.

Further, the $\log_2 S - \log_2 K$ distribution of surname occurrence (S) and its frequency K for husband, wife, and

Table 1

Frequencies of Different Surnames Among Husband and Wives (Panggi and Non-Panggi) in the Adi Panggi Tribe of Arunachal Pradesh State

Sur- Name ID	Sur- Name	Hus- bands	Wives Panggi Only	Wives Panggi and Non- Panggi	
		N=154	N=130	N=147	
1	Paron	15.58	16.92	14.97	
2	Panyang	22.08	29.23	25.85	
3	Taku	5.84	10.00	8.84	
4	Panggeng	8.44	3.85	3.40	
5	Tagi	7.79	3.85	3.40	
6	Mongku	10.39	3.08	2.72	
7	Tosang	2.60	2.31	2.04	
8	Jopir	0.65	0.00	0.00	
9	Taying	1.95	2.31	2.04	
10	Tatung	1.30	1.54	1.36	
11	Ejing	8.44	3.85	3.40	
12	Paloh	3.90	3.08	2.72	
13	Padun	1.30	5.38	4.76	
14	Tateh	3.90	3.08	2.72	
15	Tayom	1.30	1.54	1.36	
16	Aje	1.95	0.00	0.00	
17	Tarang	1.95	2.31	2.04	
18	Taruk	0.65	3.85	3.40	
19	Teksin	-	0.77	0.68	
20	Kirom	-	0.77	0.68	
21	Gete	-	1.54	1.36	
22	Tangu	-	0.77	0.68	
23	Non- Panggi (15)	-	-	11.56	

for the total sample, was considered separately in the present study. This distribution is expected to show a linear trend under the assumptions of the neutral allele model (Kimura, 1983). We also have considered the Pareto distribution (Fox 1983), where the relationship between the logarithms of n and k is expected to be

Table 2

		GT	PR	RK	KM	SS	SB
Husband	v	0.1144	0.1606	0.0966	0.0560	0.1083	0.2750
	α	4.2612	4.4005	1.2832	1.0085	4.9796	10.6207
Wife (Panggi)	v	0.0967	0.3444	0.6965	0.1467	0.1334	0.3131
	α	2.5678	9.4574	25.2379	2.5790	5.3877	11.7380
Wife (Panggi and	v	0.1679	0.4706	0.7450	0.1827	0.1607	0.3444
Non-Panggi)	α	5.6482	21.3309	40.8934	3.5767	7.0828	14.7090

Indirect Estimation of Migration of Surnames (v) and Surname Diversity (α) in Different Villages of the Adi Panggi Sub-Tribe

linear, where n is the number of individuals and k is the number of surnames, but the results are presented for the log-log distribution only.

Surname distribution and progeny size

Occurrence of surname with respect to variation in progeny size distribution (separately for male, female, and total children) among the post-reproductive women and their husbands were analyzed to investigate the pattern of surname distribution in a system of patrilocal marriage, where males (kin group, especially brothers) reside in the same village and females move out of the village after their marriage. Similar analysis was carried out for the case of reproductive wives and their husbands' sample.

Results

The occurrence of different surnames among husbands and wives of the Adi Panggi tribe are shown in Table 1. There are 18 different surnames among 154 husbands and 22 surnames among 130 wives among the Panggi tribe distributed over six villages. There are 15 surnames of non-Panggi origin that have filtered into the population through marriages with non-Panggi wives. Husbands represent a smaller number of surnames whereas wives have more diversity of surnames. On average, a single surname is shared by about 8.5% of husbands, while in the case of wives it is about 6%. The three most common surnames among husbands include Paron, Panyang, and Mongku, which represent about 47% of the husbands. In case of wives, the surnames Panyang, Paron, and Taku are the three most common and occur among 57% of Panggi wives. These seven surnames (which have a frequency above 5%) represent 78.5% of the husbands and 61.5% of thei wives. The most frequent surnames occur among both husbands and wives with similar percentages, and the seven most frequent (above 5% occurrence) constitute 70.7% of the sample. The sample contains four singly occurring Panggi surnames, representing 1.3% percent of the total individuals. Considering both Panggi and non-Panggi surnames, there are 18 singletons that constitute 6% of

	Sample Size	v	α	V _F *	α _F *
Husband	154	0.0344	5.4830	0.0513	8.3278
Wife (Panggi)	130	0.0555	7.6390	0.0483	6.6020
Wife (Panggi and Non-Panggi)	147	0.9010	14.7215	0.0559	8.7113

Table 3Estimates of Migration and Surname Diversity in Adi Panggi

* v_F and α_F are distributions based upon random isonymy.



A) Distribution of surnames among Adi Panggi Husbands

Log₂ S



B) Distribution of Surnames among Adi Panggi wives

Log ₂ S

Figure 2. Log-log distribution of surnames in Adi Panggi husbands and wives.

the sample. The unbiased random isonymy I_{ij} for the Adi Panggi samples is 0.1025.

Table 2 shows the estimates \mathbf{v} and $\boldsymbol{\alpha}$, the two parameters of Karlin-MacGregor and Fishers' distributions that describe the differential migration rates and surname diversity, among six villages of Adi Panggi. Both the husbands and wives show wide variation in the rate of migration and surname diversity between the villages. The two villages KM and RK show the least migration rates and surname diversity, whereas SB shows the highest values of \mathbf{v} and $\boldsymbol{\alpha}$ in case of the husband samples.

Panggi wives show higher surname diversity and migration rates than husbands, except in case of GT. Panggi and non-Panggi wives show higher values than husbands (Table 3). The estimates of v_F and a_F , based on random isonymy, show lesser values than from the Karlin-MacGregor and Fishers' distributions. The results show a similar pattern among husbands ($v_F = 0.0513$ and α_F = 8.3278) and Panggi plus non-Panggi wives ($v_F = 0.0559$ and $\alpha_F = 8.7113$).

The logarithmic distribution of surname occurrence among husbands and wives and in the total sample for the Panggi is shown in Figures 2 and 3.

Both the figures show a non-linear trend suggesting deviation from the expected neutral allele model. The shape of the distribution is different in husbands and wives, it is

an inverted parabolic curve in males and a slightly concave curve in females. The best non-linear fit of the curve with polynomial degree four show a variance of about 60% ($R^2 = 0.59$) in the case of the husbands sample while for wives a polynomial degree three shows a good fit to the distribution ($R^2 = 0.80$), which is better than the linear fit $(R^2 = 0.74)$. In the case of all Adi Panggi (both husbands and wives) the distributions show a nonlinear trend which is different from the trend observed in the case of husband and wife samples (Figure 3). The best fit to the curve is a polynomial of degree four, which accounts for about 96% of the variance. The non-linear distribution is due to the higher frequency of occurrence of rare surnames (especially those that occur only once and are contributed by wives). However the truncated log-log distribution (with the rare surnames excluded) shows a linear trend.

The surname distribution observed among husbands and wives has also been considered with respect to their progeny size to investigate variation of fertility differentials with respect to surnames among the post-reproductive wives and their husbands, separately for male, female, and total progeny size. The results of the distribution show a non-linear trend suggesting a differential reproductive fitness for different surnames among husbands and wives (Figure 4). In the case of husbands, the rare surnames (occurring once or twice) and common surnames (occurring more than 6-9



Distribution of surnames among the Adi Panggi tribe

Figure 3. Log-log distribution of surnames in the Adi Panggi tribe.

times) show a trend toward higher progeny size (males, females and total children) than those surnames that occur less frequently (between 3-5 times). The best fit curve in case of total children (polynomial degree three) accounts for about 94.5 percent of the variance ($R^2 = 0.945$). In case of females, the trend shows a negative association between occurrence of surnames and the progeny size and the best fit curve for the total children accounts (polynomial degree three) about 99 percent of the variance ($R^2 = 0.989$). Progeny size shows a decreas-

ing trend with the rare surnames that occur once or twice and a trend of lesser progeny in case of more common surnames (except in case the most common surnames that occur ten times among wives). These results are a reflection of the population structure of the tribe, especially the patrilocal system of marriage, which tends to increase the prevalence of some specific surnames as a result of preferential marriages among a few specific clans or surnames.





Figure 4. Occurrence of surname distribution (s); total number of male (\blacksquare), female (Δ), and total children (\times) among Adi Panggi husbands and wives (above 45 Years) and progeny size distribution (--).

imilar distribution of progeny size with respect to surname distribution was also considered in case of reproductive wives and their husbands' sample (Figure 5). The pattern of distributions is quite varied compared to the case of post-reproductive samples. In the case of the wives, the progeny size shows an increasing trend from the rare surnames to the most frequent surnames, though there is some perturbance at the initial case of rarer ones. Among the husband samples, a similar initial perturbance of progeny size is noticed with the rarer surnames, but the progeny size shows stability with the more commonly occurring surnames of the tribe.

Discussion

Though selective neutrality of surname distribution has been validated in a number of populations in Europe and Latin American countries, a few studies have shown deviation from neutrality, implicating the role of migration and operation of selection with respect to certain surnames. The results of the present study indicate that such trends are also observed in small isolated populations. Apart from the expected influences of migration and drift, there could be other related factors of population structure that could cause the deviation from neutral-





Figure 5. Occurrence of surname distribution(s): total number of male (\blacksquare), female (Δ), and total children (\times) among Adi Panggi husbands and wives (below 45 years) and progeny size distribution (--).

ity. In small tribal populations the structure of the surname distribution might vary widely depending on the history of settlement, prevailing marriage patterns and other socio-cultural and environmental reasons. Further, the demographic parameters can vary drastically owing to their subsistence pattern, epidemiological factors, and social organization such as warfare history. Also, internal tribal factions, which are formed based on kinship or clan or lineages involving several closely knit families (e.g., in South Amerindian tribes; Neel 1968; Salzano et al, 1967) may play a role. This can significantly influence the population structure and can lead to perturbations in the surname distribution in short time periods and in such cases the neutral theory predictions may not hold. Some of the above factors are important in explaining the deviation from neutrality in surname distribution among the Adi Panggi.

In general, the Adi tribe consists of several sub-tribes who occupy different terrains in the upper Himalayan mountain ranges and some in plains in the northeastern part of the country (Roy 1960; Lego 2005). The formation of the sub-tribes followed the fission-fusion model of population structure, and was triggered by intertribal factions competing for survival necessities in this hunting-gathering population (Neel 1968; Salzano et al, 1967). The formation of Adi Panggi sub-tribes occurred a couple of generations ago and was a result of conflict between a major subgroup of the tribe and some other close-kin factions over hunting rights and other resources (Roy 1960; Lego 2005; Lego 2006). The sub-tribes now occupy a recognized territory and constitute a few close affinal and consanguineous lineages and their affiliated clans. Since they have a patrilocal society, a majority of their men stay in the home village, whereas females come from different villages and from other Adi sub-tribes. This results in more female marital migration and greater surname diversity among wives. We therefore observe a higher proportion of homonymy of surname diversity among males. For example, the five most frequent surnames represent about 65% of total husbands, while rare surnames (that occur either once or twice) constitute about 5 per cent. The three most common surnames among wives also represent a large part of the population, but in contrast with the males, the rare surnames that occur either once or twice constitute 18 per cent, more than 3.5 times the proportion of rare surnames in husbands. In the total sample of husbands and wives, six surnames represent 69 percent of the total population studied. The surname distribution therefore reflects greater prevalence of most frequent (occurring in 20 to 70 cases) and rare surnames (occurring once or twice) as against the moderately frequent surnames (that occur 3 to 20 times in the sample).

The estimates of unbiased Isonymy also differ between husband and wives samples. The combined samples (husband and wives) show higher values of estimated Isonymy (I = 0.1025), suggesting higher levels of homozygosity as a result of more frequent marriages contracted

among a few selected surnames, generally found in such isolated tribes. This is reflected in the Karlin and MacGregor distribution, which shows the non-linear trend, and the trend remains the same even after the logarithmic transformation. This also explains the wide differences between the male and female surname distributions and the deviations observed in the Karlin and MacGregor and logarithmic distributions. Therefore, if the truncated distribution (omitting the rare surnames) is considered, it shows a linear trend in the combined sample, suggesting that the non-linearity is contributed by the occurrence of rare surnames in the sample. The rare surnames are contributed more through females migrating from other places. However, the discrete Pareto distribution shows a linear trend for both the samples. The linear trend is clearer in the female sample than among males, but the genetic implications of the Pareto distributions are not apparent (Fox 1983).

The results of the distribution of surnames also show wide differences with respect to progeny size distribution in case of post-reproductive (completed fertility) and reproductive wives and their husbands. The neutral allele model would suggest that the distribution should be independent of the fertility and morality (Manrubia 2002), but the present study indicates otherwise and there appears to be some differential fertility for some surnames. The Adi Panggi sample shows greater progeny size (male, female and total surviving children) for the most frequent and rare surnames, and that is reflected in the near U-shaped distribution especially in case of husband sample. This suggests reproductive advantage, especially greater success for the most frequent (6-9/10)and rare surnames (1-2) as against the less frequent surnames (3-6). Among the Adi Panggi, the four most frequent surnames among husbands and wives (especially Panyang, Paron, Mongku and Taku) show higher progeny size among the husband sample, whereas the wives sample shows a near linear trend of decreasing progeny size from the rare surnames to the most frequent ones. A paired t-test between (eight) different types of surnames and their progeny size shows higher significance values in case of the husbands (above 45 years) sample (p =(0.000433) than among the wives (above 45 years) (p = 0.003263). The mothers and their husbands were all in post-reproductive period and had attained their maximum fertility potential and do not use family planning and therefore, the possibility of other intervening factors can be excluded. It is possible that it could be also due to selective mortality of children, especially higher mortality in case of less frequent surnames than with more and rare frequent ones, and this needs to be verified. However there are no apparent cultural active selective mechanisms that can influence the progeny size among some families of the tribe. A detailed study on the marriage pattern and surname or clan distribution might reveal more details. The progeny distribution among women of completed fertility is expected to follow Poisson and Negative binomial distributions (Cavalli-Sforza

In the Indian context, surname analysis has been carried out in a few populations for investigating the effect on population structure of inbreeding (Bhalla 1976; Kashyap 1980; Vasulu 1980), genetic kinship (Kashyap 1980; Vasulu 1980), preferential mating (Vasulu 1989; Reddy 1995; Maji et al, 2007) and social network analysis among castes and tribes (Rao 1995). Surname neutrality has also been validated in a caste community of Indians residing in Texas (Chakraborty 1990).

The above few studies suggest that isonymy can be of much value in inferring population structure in some castes and tribes (Kashyap 1980; Maji et al, 2007). The surname structure in India varies among different regions, castes, tribes, linguistic and religious groups, and provides an opportunity to study microevolution and population structure where the assumptions of surname or isonymy analysis are valid. Its applicability in diverse populations, especially among isolated populations is expected to yield further insights of its validity and inferring population structure and microevolutionary history.

The present study illustrates the importance of surname in inferring the genetic structure among a small sub-tribe Adi Panggi, an isolated Tibeto-Burman- speaking community in Arunachal Pradesh state in the northeastern region of India. The tribe lives on shifting-hunting-andgathering subsistence and is a patrilocal and patrilineal society where the marriage pattern involves movement of wives from other villages to the husband's village, which allows more surname diversity through females. As a consequence of differential surname structure between husbands and wives, the surname distribution shows deviation from the expected neutral allele model and it also exhibits a non-linear trend of progeny size distribution with respect to surname occurrence in the husband sample.

The results of the present study illustrate that surname distribution could be of much value in deciphering the population structure and microevolution of a population (Crow 1965; Lasker 1978; Yasuda et al, 1974). The non-linear trend observed in this study appears to be unique in the Indian context, especially for the small isolated tribes.

The results of the study will help in further confirming the validity of the findings in several other tribes in India. In the absence of any biological or genetic data for a tribe, the surnames provide insights into the population structure and possible genetic implications which can help in the design of future studies. As a prelude to the molecular genetic investigations about the past genetic history, a surname analysis can provide considerable insight into the expected genetic structure of the population. It also helps us to frame specific hypothesis related to male and female migration history, which we will be testing with the molecular genetic studies. The two approaches together can help to unravel the differential paternal and maternal past genetic history of the tribe.

Acknowledgements

We thank the Government of Arunachal Pradesh and all the participants for their active cooperation in our study. Our sincere thanks go to Mr. Tozing Tagi and Mr. Oyam Panyang (Katan and Geku circles), for their assistance in data collection. We acknowledge the support and intellectual inputs about the Adi tribes from Mr. R. N. Koley (ARO, Directorate of Research) and Mr. Kaling Borang of Pasighat. We thank Ms. S. Krithika, Senior Research Fellow, Biological Anthropology Unit, for suggestions and help with the analysis and draft finalization, and the Director, Indian Statistical Institute, Kolkata, for the financial assistance for the study among the tribes of Arunachal Pradesh.

References

Barrai I, Barbujani G, Beretta M, Maestri I, Russo A, Formica G, Pinto-Cisternas (1987) Surnames in Ferrara: distribution, isonymy, and levels of inbreeding. *Ann Hum Biol*, 14:415-423.

Barrai I, Formica G, Barale R, Beretta M (1989) Isonymy and migration distance. *Ann Hum Genet*, 53:249-262.

Barrai I, Formica G, Barale R, Scapoli C, Canella R, Beretta M (1990) Isonymy in emigrants from Ferrara in 1981-1988. Ann Hum Biol 17:7-18.

Barrai I, Scapoli C, Formica G, et al (1991) Isonymy in records of births and deaths in Ferrara. <u>Ann Hum Biol</u>, 8:395-404.

Barrai I, Formica G, Scapoli C, et al (1992) Microevolution in Ferrara: Isonymy 1890-1990. *Ann Hum Biol*, 19:371-385.

Barrai I, Scapoli C, Bereta M, et al (1996) Isonymy and the genetic structure of Switzerland. I. The distributions of surnames. <u>Ann Hum Biol</u> 23:431-455.

Barrai I, Rodriguez-Larralde A, Mamolini E, et al (2000) Elements of surname structure of Austria. <u>Ann Hum Biol</u>, 27:607-622.

Barrai I, Rodriguez-Larralde A, Manni F, et al (2002) Isonymy and isolation by distance in the Netherlands. *Hum Biol*, 74:263-283.

Barrai I, Rodriguez-Larralde A, Manni F, Ruggiero V, Tartari D, Scapoli C, (2003) Isolation by language and distance in Beligium. *Ann Hum Genet*, 68:1-16.

Bhalla V, Bhatia K (1976) Isonymy in a Bhatia leut. *Ann Hum Genet*, 39:479-500.

Bhattacharjee PN (1954) Blood group Investigations in the Abor Tribe. *Bull Dept Anth*, 3:51-54.

Blackburn S (2004) Memories of migration: Notes on legends and beads in Arunachal, Pradesh, India. *Eur Bull Himalayan Res*, 25/26:15–60.

Boldsen J (1992) Geographical distribution of some Danish surnames: reflection of social and natural selection. *J Biosoc Sc*, 24:505-513.

Cavelli Sfroza LL, Bodmer WF (1971) *The Genetics of Human Populations.* W. H. Freeman and Co., San Francisco.

Chakraborty R, Weiss KM, Rossmann DL, Norton SL (1981) Distribution of last names: a stochastic model for lilkelihood determination in record linking. In: Morrill W, Dyke B (ed.), *Geneological Demography.* Academic Press, New York.

Chakraborty R, Schwartz RJ (1990) Selective neutrality of surname distribution in an immigrant Indian community of Houston, Texas. *Am J Hum Biol*, 2:1-13.

Colantonio SE, Lasker GW, Kaplan BA, Fuster V (2003) Use of surname models in human population biology: A review of recent developments. *Hum Biol*, 75:785-807.

Crow JF, Manage AP (1965) Measurements of Inbreeding of the frequencies of marriages between persons of the same surnames. *Eugenics Quarterly*, 12:199-203.

Das TC (1953) Social Organization of the tribal peoples. *Ind J Soc Work*, XIV:245-262.

Dyke B, James AV, Morrill WT (1983) Estimation of random isonymy. *Ann Hum Biol*, 10:295-298.

Fox WR, Lasker GW (1983) The distribution of surname frequencies. *Int Stat Rev*, 51:81-87.

Karlin S, MacGregor J (1967) The number of mutant forms maintained in a population. *Proc Fifth Berkeley Symp Math Statist Prob*, 4:415-438.

Kashyap LK (1980) Trends of Isonymy and inbreeding among the Ahmadiyyas of Kashmir. J BioSoc Sci, 12:19-25.

Kashyap LK, Tiwari SC (1980) Kinetics of genetic kinship as inferred by isonymy among the Ahmadiyyas of Kashmir valley. *Hum Biol*, 52:311-324.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Koley RN (2005) Socio-cultural aspects of Adi Panggi. Directorate of Research, Government of Arunachal Pradesh, Itanagar (in press).

Krithika S, Maji S, Vasulu TS (2006) Genetic heterogeneity among three Adi tribes of Arunachal Pradesh, India. *Hum Biol*, 78:221-227.

Lasker GW (1978) Increments through migration to the coefficient of relationship between communities by isonymy. *Hum Biol*, 50:235-240.

Lego N (2005) History of the Adis of Arunachal Pradesh. Mrs. Ponung Lego, Itanagar.

Lego N (2006) Modern History of Arunachal Pradesh (1825 – 2006). Mrs. Ponung Lego, Itanagar.

Maji S, Krithika S, Vasulu TS (2007) Genetic kinship among an isolated Adi tribe of Arunachal Pradesh: isonymy in Adi Panggi. *Hum Biol*, 79:321-337.

Manrubia SC, Zanette DH (2002) At the boundary between biological and cultural evolution: The origin of surname distributions. *J Theor Biol*, 216:461-477.

Neel JV, Chagnon NA (1968) The demography of two tribes of primitive relatively unacculturated American Indians. *Proc Natl Acad Sc* (USA), 59:680-689.

Rao AR (1995) Reciprocity in marital and social networks: Illustration with Indian data. *Hum Biol*, 67:887-904.

Reddy BM, Malhotra KC (1995) Relevance of surname in population structure: A study among the Vadde, a fishing community of Kolleru lake in Andhra Pradesh, India. *J Hum Ecol*, 6:161-168.

Relethford JH (1988) Estimation of Kinship and Genetic distance from Surnames. *Hum Biol*, 60:475-492.

Rodriguez LA, Barrai I, Alfonzo JC (1993) Isonymy structure of four Venezuelan states. *Ann Hum Biol*, 20:131-145.

Rodriguez-Larralde A, Formica G, Scapoli C (1993) Microevolution in Perugia: Isonymy 1890-1990. <u>Ann Hum Biol</u>, 20:261-274.

Rodriguez-Larralde A, Gonzales-Martin A, Scapoli C, Barrai I (2003) The names of Spain: A study of the isonymy structure of Spain. *Am J Phy Anthrop*, 121:280-292.

Roy S (1960) Aspects of Padam Minyong Culture. Directorate of Research, Govt. of Arunachal Pradesh, Itanagar.

Roy S (1966) Anthropometry of the Adis of the Siang district. North-East Frontier Agency, Shillong.

Salzano FM, Neel JV, Maybury Lewis D (1967) Further studies on Xavante Indians I. Demographic data on two additional villages and genetic structure of the tribe. *Am J Hum Genet*, 19:463-489.

Sarma J (1960) Puberty, Marriage and Childbirth among the Pangi and the Minyong Abor women. *Anthropos*, 55:97-113.

Scapoli C, Rodriguez-Larralde A, Beretta M (1997) Correlation between isonymy parameters. *Int J Anthrop*, 12:17-37.

Singh IP (1994) *People of India, National Series, Vol. III. The Scheduled Tribes.* Anthropological Survey of India, Oxford University Press, New Delhi.

Vasulu TS (1980) *Demographic Study of a Genetic Isolate: Genetic Implications of Demography of Yanadi.* Anthrop Soc of India, Mysore.

Vasulu TS (1989) Genetic structure of a tribal population 1: Breeding isolation among the Yanadis. *Int J Anthrop*, 4:255-268.

Yasuda N, Cavalli-Sfroza LL, Skolnir M, Moroni A (1974) The evolution of surnames: an analysis of their distribution and extinction. *Theor Pop Biol*, 5:123-142.

Zei G, Guglielmino CR, Siri E, Moroni A, Cavalli-Sforza LL (1983a) Surnames as neutral alleles: Observation in Sardinia. *Hum Biol*, 55:357-365.

Zei G, Matessi RG, Siri E, et al (1983b) Surnames in Sardinia I. Fit of frequency distributions for neutral articles and genetic population structure. <u>Ann Hum Genet</u>, 47:329-352.