Gene diversity among some endogamous population of Amravati District, Maharashtra, India.

Rupali Tekade[1]*, Nandkishor Warghat[1]*, Navin Sharma[1], Ashwini Wankhade and Mumtaz Baig[1#]

1. Department of Zoology, Govt. Vidarbha Institute of Science & Humanities, Amravati-444604, Maharashtra-India. E mail: nanduwarghat@gmail.com

 * This auhtor contributed equally to this work.

Corresponding author: Dr. M.M. Baig, Dept. Of Zoology, Govt. Vidarbha institute of Science and Humanities, Amravati. Email: mumtazbaig@gmail.com

Abstract:

The present work deals with the distribution of ABO, Rh and sickling alleles as a marker to study the genetic structure and micro-demarcation among castes and tribal population from Amravati district of Maharashtra. Three loci namely, ABO, Rh and sickling were selected to measure the relative frequency of respective alleles in ten (10) endogamous population inhabiting Amravati District. The ABO locus was found to be less polymorphic when compared with other loci. On the other hand Rh and sickling loci were found to be more polymorphic. Construction of dendrogram using allele frequency data reveals interesting relationship among the caste and tribe. An analysis show three major clade comprising A, B and C. The clade A comprises of Islamic *Dawoodi Bohra* and *Hindu Gujrati.* Clade B comprising upper caste as *Brahmin, Jain, Kashmiris* and *Kunbis*, while Clade C shows *Gonds* and *Katchhi*. This study is a first attempt to provide genetic landscape of caste and tribes inhabiting Vidarbha region. The findings are discussed in the light of historical, anthropological and genetical data available for the studied group.

Key words: Diversity, Caste, Tribe, Endogamous, Vidarbha.

Introduction:

India is known for its enormous genetic diversity and rank second in the world after Africa. The Indian population provides excellent opportunity to study socio-culture and genetic variability. Due to its geographical location India has served as major corridor for the dispersal of modern humans out of Africa ~2,000000 years ago (Cann, 2001).The contemporary Indian population is classified as tribal and non-tribal. There are 5000 endogamous groups in India which are known for their endogamous units and diversity (Sing, 1999). Many population groups migrated to India from northwest and northeast. It is believed that the ancestor of some of the existing tribes in southern India entered the country from the northwest, in pre-historic times. Non-tribes also came from this direction. Tribal are the aboriginal groups who have not been absorbed in the caste system (Vidyarthi, 1985) and are mainly ancestor worshipper. Linguistically population of India is classified into four major classes' viz. Indo-Aryan, Dravidian and Austro-Asiatic and Tibeto-Burman. The Indo-Aryan forms the largest group and comprises much of the north and northwest region of the subcontinent. Great deal of literature exists on classical genetic studies carried out among Indian population (Cavalli-Sforza *et al*.1964; Papiha, 1996) demonstrating high degree of heterogeneity among different caste and tribal groups. Social restrictions in marriages have apparently hampered the gene flow between different castes thereby resulting into isolated endogamous groups. Many studies on Indian population put forwarded to explain their genetic origin and genetic relationship (Bamshad, 1998; Roychaudhary, 2001; Baig, *et al*. 2004).The genetic relationship among different caste groups are not uniform. Polymorphism are now considered useful tool for studying differentiation in human population. This is generally done by using gene frequency data of number of genetic loci. Twenty two blood group loci are found to be polymorphic in human population, out of which ABO and Rh where two such polymorphic loci. Sickle cell anemia was the first polymorphism observed in

protein coding gene by electrophoresis (Pauling *et al*. 1949; Ingram, 1957). The variation in the frequency of Rh negative gene i.e. Rh (d) reported in majority of Indian population is 15-30%. The gene frequency of Rh (d) gene is lower in the tribes of Maharashtra while the Sickle-cell trait is found mainly in the tribes. The frequency of the Sickle-cell trait varies between 5 to 30 percent in many tribes of Gujarat, Maharashtra, Madhya Pradesh and Andhra Pradesh (Kate, 2002).

The Amravati city in Maharashtra is situated between latitude $20^056'$and $20^093'$ North and $77^045'$ and $77^075'$ east longitude and its total area is 12,626 km^2. It has an average elevation of 343 meters and lies 156 km west of Nagpur. Amravati has tropical wet and dry climate with hot, dry summer from March to June. The population of district is 2,607160 (census of India 2001) comprising Hindus, Buddhist, Muslims and various tribal populations. The present study was conducted during June 2007 to April 2008, sampling was done at random and healthy unrelated individuals of different endogamous group's viz. *Brahmin, Hindu Gujrati, Kunbis, Nav-Buddhist, Gonds, Korkus, Dawoodi Bohras, Muslim Katchhi, Jain* and *Kashmiris* inhabiting Amravati district of Maharashtra were sampled. In the current study, we screened 10 endogamous populations for three important loci to infer the origin and relationship amongst them.

Materials and methods:

Five ml of blood was collected from following endogamous groups using appropriate informed consent. The blood was collected in EDTA vials at random from a total 250 healthy and unrelated individuals. Twenty five (25) samples were collected and transported to the laboratory quickly in cold storage for screening. The details of endogamous group are as follows.

Brahmin: *Brahmin* is a class of priests and preachers of '*Dharma*' and considered as the torch bearer of Hinduism. Majority of *Brahmin* in Maharashtra speak Marathi, one of the

3

major languages of Indo-Aryan linguistic group. The population of *Brahmins* in Amravati district is 21,500 or 3 percent of the population. We collected 25 samples of Brahmins living in Amravati district.

Hindu Gujrati: The *Gujarati* people live in the northwestern part of the Indian subcontinent, in the state of Gujarat many *Gujarati* also inhabit Rajasthan , Maharashtra and Madhya Pradesh of India. Additionally, the Gujarati language has been adopted by communities such as the *Kachhi* and *Parasi*. Gujarati is an umbrella term used to describe an Indo-Aryan ethnic group that is traditionally Gujarati speaking.

Kunbi: *Kunbi* was claimed to have been warriors and land lord having their roots in Maharashtra. They belong to the Hindu caste and are considered as close to the upper class *Brahmin* from the same geographical location .Other than Maharashtra Kunbi are also found in Karnataka, Gujarat, Goa and Madhya Pradesh. They speak Marathi Language (a branch of Indo-Aryan linguistic family) and make more than 30% population of Maharashtra.

Nav Buddhist: *Nav-Buddhist* is those who have adapted Buddhism some 50 years ago and hence called as *Nav-Buddha* (New- Buddhist). These are widely spread through out India and make up 16 % of Indian population. Buddhist also speaks Marathi language

Gonds: This tribe fall under primitive tribes category and spread much over the central India. *Gond* generally speaks '*Gondi*' dialect which belong to Dravidian linguistic family and is consider being closer to Tamil.

The Korkus: The *Korkus* are typical tribal population from Amravati district and found only in the *Satpuda* mountain ranges spanning Maharashtra and Madhya Pradesh. They are mainly concentrated in *Melghat* a scheduled area of Korku comprising 89 % of the tribal population. The *Korkus* speak *Korku* dialect belonging to Austro-Asiatic linguistic group. The Austro-Asiatic speakers are considered as the first settler of Indian subcontinent (Roychaudhary, 2001).

The Dawoodi Bohras: Dawoodi Bohras are the main branch of Bohras community in India. Today there approximately are one million *Dawoodi Bohras* resides in Mumbai, Surat and Madhya Pradesh. Some Bohras ancestors were converted from Hinduism to Islam. Bohras have their own language called Lisanu-I-Dawat written in Arabic script but is derived from Urdu and Gujrati.

Muslim Kachhi: Muslim Kachhi are Sunni Muslim who migrated from Sindh to Kutch in Gujarat state of India. After their conversion of Islam in 1422 C.E. Kachhi speaks non-written dialect called Memoni, a mixture of Sindhi and Kachhi which belongs of Indic family. The collected samples of Kachhi population are native of Amravati district.

Jain: Jainism present in Maharashtra since ancient times. The emigrant *Jains* has a large population in Maharashtra. Majority of them are from Rajasthan and Gujarat. Jains speak non-written dialect' Marwari' originated from Indo-European language group.

The Kashmiris: From the time of the beginning of formation of the present Asian continent, Kashmir has remained an anseparable part of Indian peninsula. The population of Jammu and Kashmir has highest proportion of Muslim (96%) and only (4%) of Hindus consisting of small minorities of Sikhs and Buddhist. The Kashmir people belong to the Aryan race and widely spoken language belonging to the Indo-Aryan branch of Indo-European language family. Some *Kashmiris* migrated in Maharashtra by the cause of terrorism.

Laboratory Analysis

ABO and Rh (D) Group Typing: Slide agglutination method was followed for ABO and Rh (D) group typing. All samples are tested for the ABO and Rh blood groups using anti – A, anti – B and anti – D sera.

Sickling Test: Sickling test was carried out by method as described by Scott and Castro (1998). Late positive sickling samples were subjected to cellulose acetate membrane electrophoresis by using method of Marengo-Rowe (1965).

5

Statistical Analysis

Measurement of Genetic variation:

Allele frequency was calculated using Hardy-Weinberg principle. Coefficients of gene differentiation (G_{ST}) and (D_{ST}) were calculated following Nei (1973). The genetic relationship among the present endogamous population group were assessed using the measure of genetic distance (D) proposed by Nei (1972). A dendrogram was drawn as per UPGMA clustering method using Phylip (v 3.69), Felsensteins (1993).

Result and Discussion

In ABO blood group, the phenotype frequency of B was found to be higher in Buddhist, Bohras, Gonds and Katchhi than the other population. The frequency of AB was higher only in the Korkus population as compare to other groups. Among total population, the frequency of B was highest (38.4) as compared to blood group O (28.0). The frequency of blood group A and AB was found to be 26.04 and 07.20 respectively (Table. 1).

Rh (D) phenotype frequency (08.00) was observed high among Bohras, Kashmiris and Gujrati population (Table. 2).The phenotype and gene frequency of Sickling was high only in the Navbuddhist population (0.40). Sickling gene was virtually absent in other population (Table. 3).

The Chi-square differences was non-significant of group Katchhi in case of ABO blood groups(X^2=0.116; df 3; p-value=0.989) and remaining population shows significant differences. In case of Rh (D) and Sickling all populations have significant value of X^2. The X^2 differences in allelic frequency for majority of markers for various population was significant (X^2=0.461, df=18, p<1.00 for ABO; X^2=0.052; df=9; p<1.00 for Rh (D) and X^2=0.058; df=9; p-value=0.999 non-significant for Sickling). The X2 values further point towards stratification of Hindu caste system into caste and sub-caste where social forces are acting as the main deterrent in the free gene flow among them.

6

Genetic diversity:

The gene diversity indices for the total population H_T and intra populational gene diversity H_S were quite high i.e. 0.926 and 0.853 (Table.4). The inter-populational gene diversity D_{ST} was high i.e.1.00, and the coefficient of genetic differentiation (G_{ST}) was 0.079. The pooled H_T, D_{ST} and G_{ST} were found to be highest for Sickling and lowest for ABO; while pooled H_S was highest for ABO and lowest for Sickling. Previous study on Indian caste and tribe also reveal higher intra-population diversity (Bamshad *et al.* 1998).

Genetic distance:

Using the allele frequency data for the five markers, Nei's measure of genetic distance was calculated between the ten (10) endogamous groups. It was found highest between Navbudhist and Kunbis (0.5792) and lowest between Gonds and Katchhi (0.0007) (Table.5). Using distance matrix a dendrogram was constructed as per UPGMA method (figure. I).The dendrogram constructed using UPGMA method clearly shows Korkus and Navbuddhist as outlier. The dendrogram shows three distinct clades designated as A, B and C. The clade A comprises of Islamic Dawoodi Bohra and Hindu Gujarati. Both Bohra and Gujrati speaks are linguistically related. The Bohra has known history of migration from Middle East and were first settle down in Gujarat (Muthukrishnan, *et al.* 2010). Another monophyletic clade B includes *Brahmin, Jain, Kashmiris and Kunbi*. Of note, the entire ethnic group in this cluster belongs to higher rank of the caste. The Kunbis are considered as sub-caste within the warrior class the "Shatriyas", more frequently referred as Maratha. The genetic similarity of *Maratha* with the *Brahmin* is well established (Cavalli-Sforza *et al.* 1994). Also Maratha are considered as the migrant group from north of India (Sahoo, *et al.* 2005). Interestingly,

Clade C shows clustering of tribal *Gond* and *Katchhi*, an Islamic group. This clustering might be due to limitation of markers employed or might be due ancient link of them. Interestingly, in the past Gonds has rule the central India (Bhasin, 1992). and Katchhi, basically a trading

7

group might have established some linkages. Taking into account the sample size, we should not read too much but some pattern are clear that signature of social force still persist in the genome of cast and tribes of India. In coming years sampling of more endogamous group and typing using DNA marker will provide better insight into the genetic landscape of Indian population.

References:

1. Baig M, Khan A and Kulkarni K. (2004) Mitochondrial DNA diversity in Tribal and Caste group of Maharashtra and its implication on their genetic origins. Annals of Human Genetics, 68: pp. 433-460.

2. Bamshad MJ, Watkins WS, Dixon ME, Torde LB, Rao BB, Naidu JM. (1998) Flow stratifies Hindus castes. Nature, 395: pp. 652-662.

3. Bhasin MK, Walter H and Dnaker-Hopfe H. (1992) The distribution of Genetic, Morphological and behavioral Traits among The People of Indian Region. Kamala-Raj Enterprises, Delhi.

4. Cann RL. (2001) Genetic clues to dispersal of human populations: retracing the past from the present. Science, 291: pp. 1742-1748.

5. Cavalli-Sforza LL and Edwards AWF. (1964) Analysis of human evolution. Proc. 11th Int. Congr. Genet. 2: pp. 923-933.

6. Cavalli-Sfroza L.L., Menozzi P. and Piazza A. (1994) The History and Geography of Human Genes. Princeton University Press, Princeton.

7. Census of India 2001. Data from the 2001 Census includes cities, villages and towns (Provisional).

8. Felsensteins J. (1993) Phylip (Phylogeny Package) version 3.69 Distributed by the author. Seattle, Dept. of Genetics, University of Washington.

9. Ingram VM. (1957) Genetic mutation in human hemoglobin. Nature (Lond.) 180: pp. 326-328.

10. Kate S.L. (2002) Epidemiology of Sickle Cell Disorder in the state of Maharashtra. Int J Hum Genet, 2(3): pp. 161-167.

11. Marenge-Rowe AJ. (1965) Rapid electrophoresis and quantitation of hemoglobin on cellulose acetate. J. Clin. Pathol. 18: pp. 790-792.

12. Muthukrishnan Eaaswarkhanth, Ikramul Haque, Zeinab Ravesh, Irene Gallego Romero, Poorlin Ramakodi Meganathan, Bhawna Dubey,Faizan Ahmed Khan, Gyaneshwer Chaubey, Toomas Kivisild, Chris Tyler-Smith, Lalji Singh, and Kumarasamy Thangaraj. (2010) Traces of sub-Saharan and Middle Eastern Lineages in Indian Muslim Populations. Fur J Hum Genet. 18(3): pp. 354–363

13. Nei M. (1972) Genetic distance between populations. Amer. Nat. 106: pp. 283-292.

14. Nei M. (1973) Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci., USA, 70: pp. 3321-3323.

15. Papiha SS, Chahal SMS, Roberts DF and Sing IP. (1996) Genetic studies among Kanet and Koli of Kinnar District in Himachal Pradesh, India. Am. J. Rhys. Anthropol. 53: pp. 275-283.

16. Pauling L, Itano HA, Singer SJ and Wells IC. (1949) Sickle-Cell anemia, a molecular disease. Science (Wash., D.C.) 110: pp. 543-548.

17. Roychaudhary S, Roy S, Basu A, Banerjee R, Vishwanathan H, Usha Rani, Sil S K, Mitra M and Mujumder PP. (2001) Genomic structure and population histories of linguistically distinct tribal groups of India. Hum Genet, 109: pp. 339-350.

18. Sahoo S and Kashyap VK. (2005) Influence of language and ancestry on genetic structure of contiguous population: A microsatellite based study on populations of Orissa. BMC Genetics. Doi: 10.1186/1471 pp. 1256-64.

19. Scott RB and Castro. (1979) Screening for sickle cell haemoglobinopathy. J. Am. Med. Ass. 241: pp. 1145-1148.

20. Singh KS. 1997. The Scheduled Tribes. In: K.S. Singh (ed.), People of India. Oxford University Press, Oxford, pp. 1266.

21. Vidyarthi LP. 1983. Tribes of India. In people of India: Some Genetical Aspects. G.V. Satyavati ed., New Delhi. pp. 85-103

Table 1. ABO blood group frequency among different endogamous population of Amravati district.

Population	Observed	Phenotypes				Allele frequency		
		A	B	AB	O	A	B	O
Navbudhist	No %	7 28.00	11 44.00	3 12.00	4 16.00	0.2373	0.3441	0.4185
Kunbi	No %	6 24.00	7 28.00	1 04.00	11 44.00	0.1623	0.1838	0.6539
Bohras	No %	5 20.00	12 48.00	1 04.00	7 28.00	0.1442	0.3146	0.5413
Korku	No %	11 44.00	9 36.00	2 08.00	3 12.00	0.3262	0.2656	0.4042
Kashmiris	No %	8 32.00	5 20.00	2 08.00	10 40.00	0.2301	0.1614	0.6085
Brahmin	No %	9 36.00	8 32.00	1 04.00	7 28.00	0.2411	0.2125	0.5464
Gond	No %	7 28.00	13 52.00	1 04.00	4 16.00	0.1867	0.3575	0.4545
Jain	No %	6 24.00	8 32.00	2 08.00	9 36.00	0.1854	0.2338	0.5808
Gujrati	No %	3 12.00	10 40.00	3 12.0	9 36.00	0.1368	0.3095	0.5536
Kachhi	No %	4 16.00	13 52.00	2 08.00	6 24.00	0.1372	0.3704	0.4845
Total (n=250)	No %	66 26.04	96 38.40	18 07.20	70 28.00	0.1368	0.1614	0.4042

Table 2. Allele frequency of Gene Rh (D) factor among different endogamous population of Amravati district.

Population	Observed	Phenotype		Allele frequency	
		Rh + (D)	Rh - (d)	D	d
Navbudhist	No %	24 96.00	1 04.00	0.7774	0.2226
Kunbi	No %	24 96.00	1 04.00	0.7774	0.2226
Bohras	No %	23 92.00	2 08.00	0.7097	0.2903
Korku	No %	24 96.00	1 04.00	0.7774	0.2226
Kashmiris	No %	23 92.00	2 08.00	0.7097	0.2903
Brahmin	No %	24 96.00	1 04.00	0.7774	0.2226
Gond	No %	24 96.00	1 04.00	0.7774	0.2226
Jain	No %	24 96.00	1 04.00	0.7774	0.2226
Gujrati	No %	23 92.00	2 08.00	0.7097	0.2903
Kachhi	No %	24 96.00	1 04.00	0.7774	0.2226
Total (n=250)	No %	237 94.80	13 05.20	0.7097	0.2226

Table 3. Sickling test, allele frequency among different endogamous population of Amravati district.

Population	Observed	Phenotype		Allele frequency	
		Sickling test +ve	Sickling test -ve	S	s
Navbudhist	No %	1 04.00	24 96.00	0.7774	0.2226
Kunbi	No %	0 0	25 100	0.8862	0.1138
Bohras	No %	0 0	25 100	0.8862	0.1138
Korku	No %	0 0	25 100	0.8862	0.1138
Kashmiris	No %	0 0	25 100	0.8862	0.1138
Brahmin	No %	0 0	25 100	0.8862	0.1138
Gond	No %	0 0	25 100	0.8862	0.1138
Jain	No %	0 0	25 100	0.8862	0.1138
Gujrati	No %	0 0	25 100	0.8862	0.1138
Katchhi	No %	0 0	25 100	0.8862	0.1138
Total (n=250)	No %	1 0.40	249 99.60	0.7774	0.1138

Table 4. Gene diversity analysis in Ten (10) endogamous population of Amravati district.

Locus	D_{ST}	H_T	H_S	G_{ST}
ABO	1.000	0.957	0.914	0.045
Rh (D)	1.000	0.922	0.845	0.084
Sickling	1.000	0.901	0.801	0.110
Pooled	1.000	0.926	0.853	0.079

Table 5. Genetic distance matrix among different endogamous population of Amravati district.

| Population	Navbudhist	Kunbi	Bohras	Korku	Kashmiris	Brahmin	Gond	Jain	Gujrati	Katchhi
Navbudhist		0.379215	0.349479	0.347934	0.357351	0.361301	0.354282	0.366483	0.351254	0.358652
Kunbi			0.007881	0.019998	0.003186	0.003725	0.014616	0.001659	0.007154	0.013307
Bohras				0.013912	0.007629	0.006183	0.004399	0.003915	0.000049	0.003289
Korku					0.015411	0.006456	0.006595	0.011036	0.015161	0.011435
Kashmiris						0.003357	0.015840	0.003608	0.007256	0.016530
Brahmin							0.006936	0.000989	0.006300	0.008443
Gond								0.006600	0.005144	0.000735
Jain									0.003733	0.006401
Gujrati										0.003753
Katchhi										
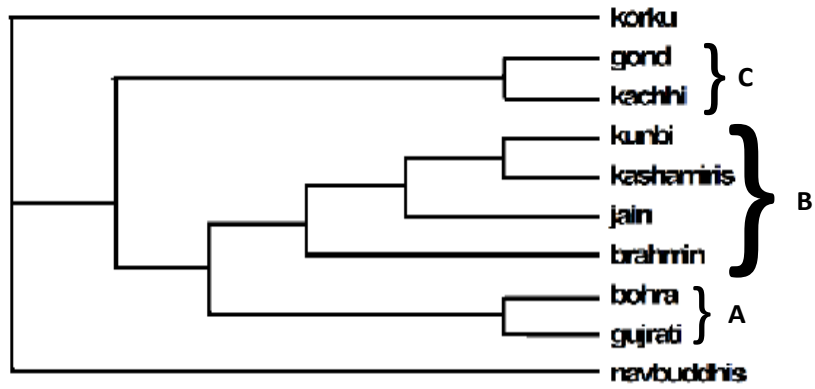
Figure I. Dendrogram showing genetic relationship among ten (10) endogamous population of Amravati district, Maharashtra.