Biotechnology International, Vol 2(4): 91-115. Dec 2009

Are Austro Asiatic tribes original native inhabitant of India?

Fahim Ahmad*, Amna Siddiqui ,Aswin Ashok Raut** and P.K. Pandey****

*Department of Biomedical Sciences and centre for excellence for infectious diseases, Paul Foster School of Medicine, Texas Tech university, 5001 El Paso Texas(USA). TX-79905; **Department of Biotechnology, Bundelkhand University, Jhansi(India).

Corresponding Author:- Dr. Fahim Ahmad

Department of Biomedical Sciences and centre for excellence for infectious diseases, Paul Foster School of Medicine, Texas Tech university, 5001 El Paso Texas(USA). TX-79905.

Key words: Homo sapiens, Austro Asiatic, mitochondrial DNA

Summary: The most current issue is about the people who had settled first in Indian sub continent. It has been suggested that the Austro Asiatic linguistic family are perhaps the first to settle in India and the Palaeoanthropological evidences suggest the earliest settlement probably around 60,000 years BC. Homo sapiens evolved in Africa and replaced archaic human in other parts in world. One of the first waves of out of Africa migration came into India. India served as a major corridor for dispersal of modern humans. Mitochondrial DNA sequence data of hypervariable segment (HVS-1) and Y-chromosomal haplogroup data indicate that the Austro Asiatic tribal population may have been the most ancient inhabitant of India.

Introduction

India has served as major corridor for the dispersal of modern human (Cann, 2001). The entry date of modern human remains uncertain, by the middle Paleolithic period (50,000 to 20,000 years before). The modern human migration routes remains enigmatic, whether there were also return to Africa from India /Asia is unclear (Maca-Meyer et al 2001, Roychodhury et al 2001, Cruciani et al 2002).

The Indian people is culturally Stratified as tribal, who constitute 8.08% of the total population (1991 census of India). There approximately 450 tribal communities in India (Singh et al 1992), who speak approximately 750 dialects (Kosambi 1991) that classified into one of three languages family. **Austro Asiatic (AA), Dravidian (DR) and Tibeto Burman (TB).** The tribal are possibly the original inhabitant of India (Thapar 1966, Roy 1973) although their evolutionary histories and biological contribution to the non tribal population have been debated (Risley1915, Guha 1935, Sarkar1958).

The Austro Asiatic (a subfamily of the Austric Language family) speaking tribes are the original inhabitant of India. Some scholars (Buxton 1925, Sarkar 1958) have, however, proposed that the Dravidian are the original inhabitants of India, the Austro Asiatic being later immigrants. The Austro Asiatic family is a fragmented languages group. It is most likely spoken in Vietnam and Cambodia. In India very small number of ethnic group, speak Austro Asiatic languages. The Indian Austro Asiatic speakers are mainly tribes, which may be indicative of the oldest inhabitant of India (Pattanayak 1998, Gadgil et al 1998). Some people believe that the Austro Asiatic family evolved in southern China (Diamond 1997).

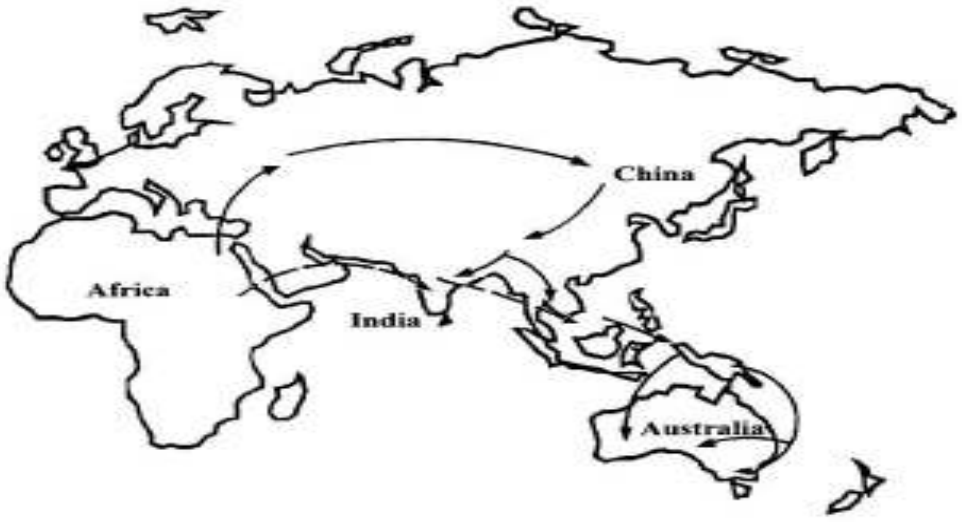


Fig.1. Schemic representation of the two alternative hypotheses suggesting possible routes of earliest migration of a people into the Indian subcontinent.

The routes by which the first migrants entered India are also not clearly understood. Basically, there are two major schools of thought on this and fig 1 depicts the plausible scenarios of the earliest migration into India. According to some scholars (Ballinger *et al* 1992; Gadgil *et al* 1998; Diamond 1997) the Austrics had their origin in China, entered India through northeast corridor and then passed onto islands beyond. A strong support for this theory comes from the fact that almost all the Austro- Asiatic tribes are located in eastern and north-eastern, central India. Archaeological evidences provided by Lal (1956) Mohapatra (1975, 1985) and Gupta (1979) support this theory.

The second principal component, derived from the analysis of 69 genes from 42 populations of Asia, which explains 17.7% of the variation is also compatible with fanning out of people from Southeast Asia and China into India (Cavalli-Sforza *et al* 1994). A more recent view is that the Austro-Asiatic speakers were another wave of migration from Africa to India and then to Southeast Asia (Nei and Ota 1991; Chu *et al* 1998; Su *et al* 1999; Majumder 2001). The skull and the general anatomy of the fossil specimen, near Panchmari, in the village of Langhnaj in Gujrat, India, were similar to the specimens found in the Northeast Africa (Kennedy 2000). Nei and Ota (1991) also suggest that Mundari groups (branch of Austro-Asiatic linguistic family) in India, which are known to have dark skin, constitute perhaps another wave of migration from Africa. The Australoid- looking skeletons have also been found in Iran and Mesopotamia (Sergent 1997). All these related evidences may suggest that the probable migratory route of Austro-Asiatic people was from Africa to the Indian subcontinent and then to Indonesia and Australia. An early wave of migration into India, actually from Africa through India, to southeast Asia has also been proposed using nuclear DNA microsatellite markers (Chu *et al* 1998) and Y-chromosomal DNA markers (Su *et al* 1999). This view is reinforced by the fact that the 9bp deletion, which was hypothesized to have arisen in Central China and radiated out from this region to southeast Asia (Ballinger *et al* 1992), is absent in most Indian populations and present in low frequency in southeast Asia (Majumder 2001). However, this inference seems to have limited value as many recent studies indicate that a number of south Indian caste and tribal populations showed 9bp deletion in variable frequency ranging up to 50% (Watkins *et al* 1999; Clark *et al* 2000; Reddy B M, Naidu V M, Madhavi V K, Thangaraj K and Singh L,

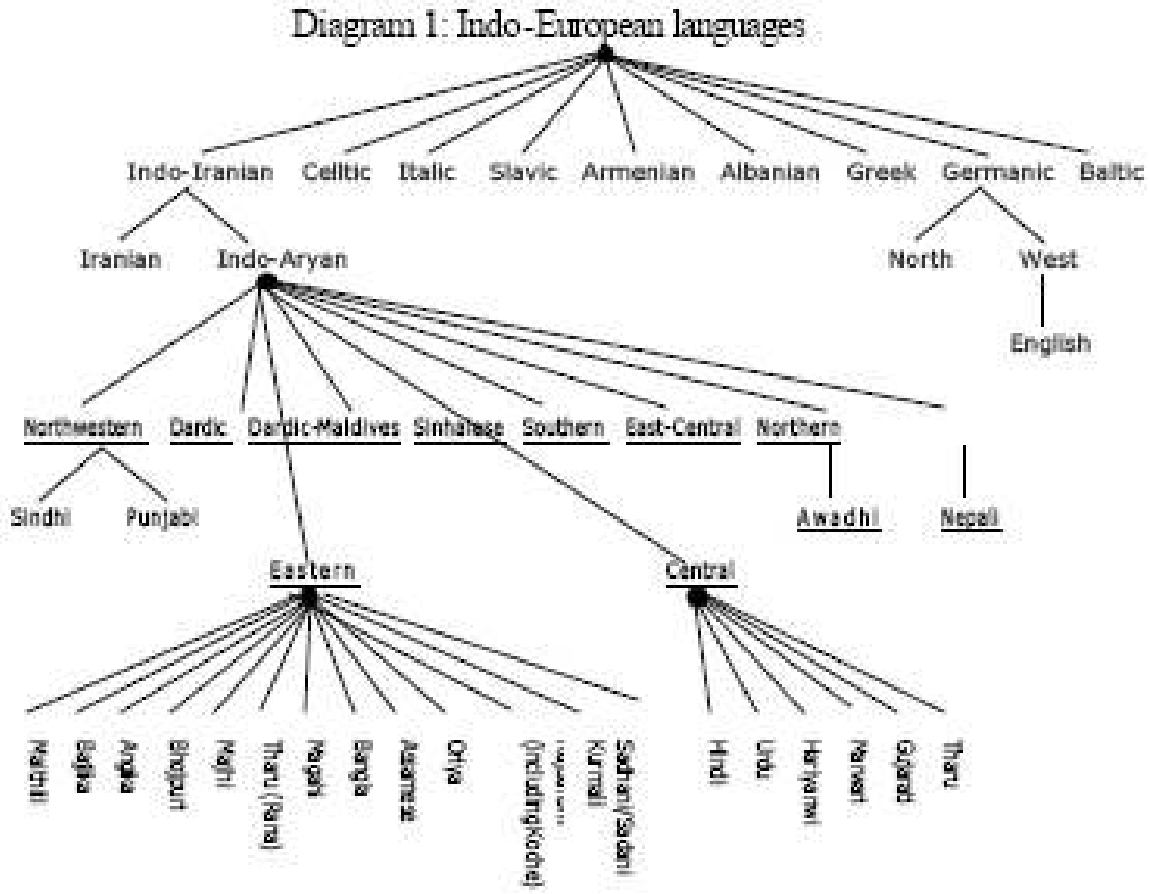
unpublished results). On the basis of high frequency of haplogroup M, Quintana-Murci *et al* (1999) have proposed that this haplogroup has originated in East Africa approximately 60,000 years BP and that it was carried into India through an East African exit route by an early dispersal event of modern humans out of Africa. Concurrent to this, the first principal component based on 69 genes from 42 populations of Asia (Cavalli-sforza *et al* 1994), which explains about 35% of the total variation in the gene frequency, suggests movement of people radially fanning from Middle East. The fact that Mundari speaking tribes inhabit mostly eastern, central and north-eastern region of India has been considered as an indirect evidence for the first theory of migration. This can be, however countered by hypothesizing the assimilation of some of the Austro-Asiatic tribes by later migrants; the disinterest of the later migrants in accessing the difficult terrain of hills and forests of central and eastern India had probably resulted in the survival of the Mundari and the other Austro-Asiatic groups inhabiting such tracts. Further, Sergent (1997) noted that the Austro-Asiatic groups were once predominant in a continuous belt, from central India to Vietnam.

Austro-Asiatic speakers in India are represented by almost 30 different tribal groups, covering a wide geographic area. Although having broad linguistic and cultural similarity, they do represent two different subgroups of languages – Mundari and Mon-Khmer – and physically belong to two racial types, the Mongoloids and Australoids. Therefore, before inferring on the origin, migration and/or antiquity of these people, it is necessary to first ascertain if all the Austro-Asiatic groups in India genetically constitute a single entity or are they derived from different sources. It is in this context that we have taken up a major endeavor to study the extent of genomic diversity within this broad linguistic category, covering almost all the major groups. As a first step towards this endeavor we gathered all the available geographic, cultural and biological data and analysed critically to examine the nature and extent of heterogeneity among the Austro-Asiatic populations the results of which are presented in the subsequent sections.

Fig. 2 Map of India, showing the geographical distribution of Austro Asiatic tribes in India.

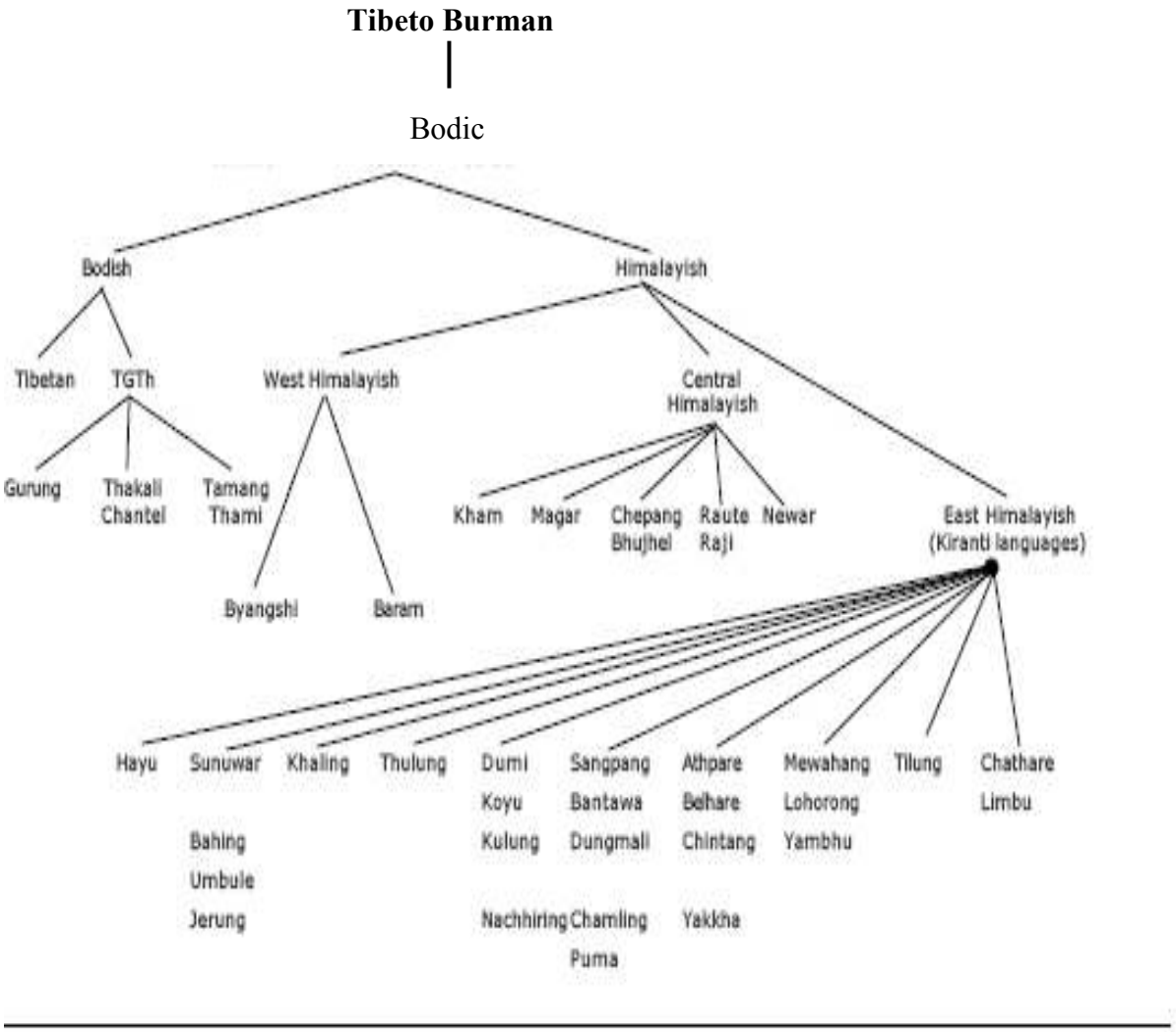
Genetic Affiliation

The languages enumerated in the 2001 census belong to the four language families. **Indo-European, Tebeto Burman**, **Dravidian,** and **Austro-Asiatic**. Indo-European family of languages mainly comprise Indo-Aryan group of languages, which forms the largest group of languages in terms of speakers nearly 80 per cent. The Indo-Aryan languages spoken in India can be genetically subcategorized in the following diagram:



Another important group of India's languages is the Tibeto-Burman group of Sino-Tibetan family. Though it is spoken by relatively lesser number of people than the Indo-European family, it consists of the largest number of languages, viz. about 57 languages. Contrary to speakers of Indo-Aryan languages, there has been steady decline in speakers of Sino-Tibetan languages from the 1952/54 to the 1981 censuses, viz. 21.8 (1952/54), 19.26 (1961), 17.16 (1971), and 12.06 (1981) per cent. In the last two censuses, they have, however, increased to 16.76 per cent (1991) and nearly 19 per cent (2001). Their decline and increase may also be due to the reasons ascribed to those of Indo-European languages.

Diagram 2: Sino-Tibetan languages



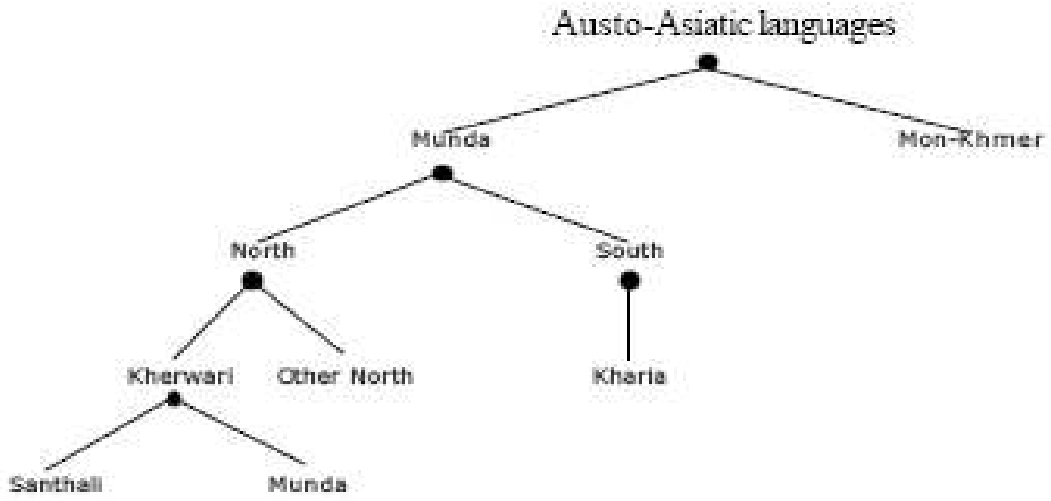
Adapted from Bradeley (2002) and David Watters (personal communication)

In addition to these two major language families, there also exist a few languages belonging to two language families. They are Austric branch of the Austro-Asiatic family and Dravidian family of languages. The Austric languages comprise Santhali of the northern Munda group and Kharia of the southern Munda group. It is to be noted that Satar has been reported in all the censuses but Santhal has been wrongly reported as a separate language except in the 1952/54 census. The 2001 census lumps both Satar and Santhal together into a single language, called Santhali. It is suggested that Munda (with 67

speakers) should also be included within Santhali, in that it is just a variant name of the same language.

According to the 2001 census, Santhali speakers are 40, 193 in number, i.e. 0.18% of India's total population, as compared to 0.20% (1952/54), 0.31% (1961), 0.21% (1971), 0.19 (1981), and 0.18% (1991). Another Austric language of Munda branch is Kharia, which has been introduced in the 2001 census for the first time. This language is spoken by 1575, i.e. 0.01%. All the Austric languages are spoken by groups of tribal peoples from the eastern Terai and make up approximately 0.19 per cent of the total population. The genetic affiliation of the Austric languages spoken in India is shown in the following diagram:

Diagram -3 Austro Asiatic languages



Robert Caldwell (1856, 3rd edn, repr. 1956: 3–6) was the first to use 'Dravidian' as a generic name of the major language family, spoken in the Indian subcontinent. The new name was an adaptation of a Sanskrit term*dravi . da*- (adj *dr¯avi.da*-) which was traditionally used to designate the Tamil language and people, in some contexts, and in others, vaguely the south Indian peoples.

Dravidians: prehistory and culture

Prehistory

It is clear that 'Aryan' and 'Dravidian' are not racial terms. A distinguished authority on the statistical correlation between human genes and languages, Cavalli-Sforza (2000), refuting the existence of racial homogeneity is a genetic homogeneity that is never achieved in populations higher animals) would require at least twenty generations of 'inbreeding' (e.g. by brother– sister or parent–children mating repeated many times) . . . we can be sure that such an entire inbreeding process has never been attempted in our history with a few minor and partial exceptions. (13) There is some indirect evidence that modern human language reached its current state of development between 50,000 and 150,000 years ago . Beginning perhaps 60,000 or 70,000 years ago, modern humans began to migrate from Africa, eventually reaching the farthest habitable corners of the globe such as Tierra del Fuego, Tasmania, the Coast of the Arctic Ocean, and finally Greenland. Calculations based on the amount of genetic variation observed today suggests that the population would have been about 50,000 in the Paleolithic period, just before expansion out of Africa and that the genetic tree and the linguistic tree have many 'impressive similarities' (see Cavalli-Sforza 2000: fig. 12, p.144). The figure, in effect, supports the Nostratic Macro-family, which is not established on firm comparative evidence (Campbell 1998, 1999). The center of origin of Dravidian languages is likely to be somewhere in the western half of India. It could be also in the South Caspian (the first PC center), or in the northern Indian center indicated by the Fourth PC. This language family is found in northern India only in scattered pockets, and in one population (Brahui) in western Pakistan suggesting a relationship between Dravidian and Elamite to the west and also the language of the Indus civilization , following the speculative discussions in the field. Still there is no archeological or linguistic evidence to show actually when the people who spoke the Dravidian languages entered India. But we know that they were already in northwest India by the time the Rigvedic Aryans entered India by the fifteenth century BC. In an earlier publication, Cavalli-Sforza et al. (1994: 239) have given a genetic tree of twenty-eight South Asian populations including the Dravidian-speaking ones, which is reproduced below as figure.

They say: A sub cluster is formed by three Dravidian-speaking groups (one northern and two central Dravidian groups, C1 and C2) and the Austro-Asiatic speakers, the Munda. The C1 Dravidian group includes the Chenchu–Reddi (25,000), the Konda (16,000), the Koya and others, all found in many central and central-eastern states, though most data come from one or a few locations. The C2 Dravidian group includes the Kolami–Naiki (67,000), the Parji (44,000) and others; they are located centrally, a little more to the west. North Dravidian speakers are the Oraon (23 million), who overlap geographically with some of the above groups and are located in a more easterly and northerly direction. (239) The second major cluster, B, contains a minor subcluster B1 formed by Sinhalese, Lambada, and South Dravidian speakers . The South Dravidian group includes a number of small tribes like Irula (5,300) in several southern states but especially Madras, the Izhava in Kerala, the Kurumba (8,000) in Madras, the Nayar in Kerala, the Toda (765), and the Kota (860in 1971) in the Nilgiri Hills in Madras (Saha et al. 1976). (240)

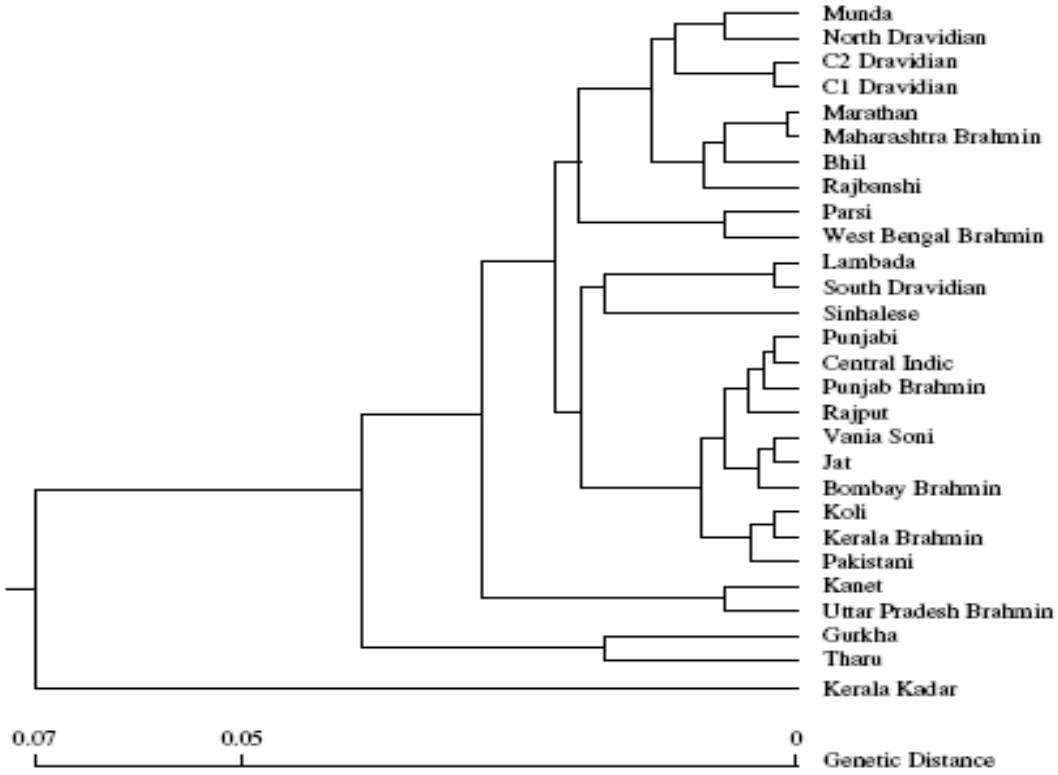


Fig - Genetic tree of South Asian populations including the Dravidian-speaking ones

Several scholars have maintained, without definite proof, that Dravidians entered India from the northwest over two millennia before the Aryans arrived there around 1500 BCE. Rasmus Rask 'was the first to suggest that the Dravidian languages were probably "Scythian", broadly representing "barbarous tribes that inhabited the northern parts of Asia and Europe" ' (Caldwell 1956: 61–2). There have been many studies genetically relating the Dravidian family with several languages outside India (see for a review of earlier literature, Krishnamurti 1969b: 326–9, 1985: 25), Revising his earlier claim (1972b) that Dravidians entered India from the northwest around 3500 BC, Zvelebil (1990a: 123) concludes: 'All this is still in the nature of speculation. Atruly convincing hypothesis has not even been formulated yet.' Most of the proposals that the Proto-Dravidians entered the subcontinent from outside are based on the notion that Brahuiwas the result of the first split of Proto-Dravidian and that the Indus civilization was most likely to be Dravidian. There is not a shred of concrete evidence to credit Brahui with any archaic features of Proto-Dravidian. The most archaic features of Dravidian in phonology and morphology are still found in the southern languages, namely Early Tamil *aytam*, the phoneme *.z*, the dental-alveolar-retroflex contrast in the stop series, lack of voice contrast among the stops, a verbal paradigm incorporating tense and transitivity etc. The Indus seals have not been deciphered as yet. For the time being, it is best to consider Dravidians to be the natives of the Indian subcontinent who were scattered throughout the country by the time the Aryans entered India around 1500 BCE.

Geographic affinity of Austro Asiatic tribes

Figure 2 presents the nature of distribution of Austro- Asiatic speakers in India. The Mon-Khmer group of Austro-Asiatic people includes the Khasi tribes, which are confined to the Khasi and Jantia hills of Meghalaya, and Nicobarese and Shompen of the Nicobar inlands. The Austro-Asiatic tribal communities affiliated to the Mundari branch are mainly concentrated in and around Chota-Nagpur plateau. While the Southern Mundari groups such as Juang, Gata, Bondo, Bodo Gadaba, Paranga and Saora inhabit the Koraput and adjoining districts of Orissa, Kherwarian groups comprising of Asur, Birhor, Ho, Korwa, Santhal, Turi and Munda are widely distributed from Jharkhand (Ranchi, Gumla, Lohardaga and Singhbhum districts) to Orissa (Mayurbhanj, Keonjhar and Sundergarh

districts) to Madhya Pradesh (Raigarh and Jashpur districts) and to West Bengal (Birbhum, Nadia and Bakura districts). A section of Korku is the only tribe that inhabits the north-eastern border areas of Maharastra. From the distributional pattern one can infer that these communities bear a strong affinity in that they occupy contiguous patches of ecologically similar habitats characterized by hilly and heavily forested areas.

Cultural homogeneity of Austro Asiatic tribes

All these communities are exclusively tribes, which are further divided into endogamous sub-tribes. Each of these tribes and subtribes constitutes within them a number of exogamous clans that are generally totemic i.e. named after plants, animals, birds and insects. These are basically hunting and food-gathering societies, supplemented by shifting cultivation. These tribes are also unique in practicing primitive technologies (Gadgil *et al* 1998). Their political organization is found to be very traditional, constituting a very simple structure. One of the cultural traits exclusively found in Mundari groups is the use of vermilion by the women of these tribes to indicate their marital status (S Barua, personal communication). From the foregoing account, it appears probable that these Austro-Asiatic tribal communities, particularly the Mundari speakers, are a homogenous cultural entity.

Ethnic affiliation of Austro Asiatic tribes

With exceptions to the Mon-Khmer groups, which are classified as Mongoloid, ethnically these tribal communities belong to Proto-Australoid type that is believed to be the basic element in the Indian population (Thapar 1996). Many scholars have classified these tribal communities under different headings. Lapicaque (1905) used the term pre-Dravidian while Chandra (1916) favoured the term Nishada. Hutton (1933) used the term proto-Australoid and even put the Veddas of Sri Lanka within this ambit. While Hooton (1930) replaced the term proto-Australoid as pseudo-Australoid, Guha (1937) used the term proto-Australoid to designate these tribal groups. Taking into account the different classifications proposed, it is clear that all these authors agree that these tribal communities belong to one ethnic group although there is no agreement on the names used for the classification. A comprehensive picture will probably emerge only when we look at biological similarities shared by Austro-Asiatic speaking tribal communities.

Biological affinities of Austro Asiatic tribes

Recently published data pertaining to molecular genetic markers are very patchy and quite inadequate to unequivocally conclude that the Austro-Asiatic tribal groups belong to a single and genetically homogeneous ethnic group. Although data on anthropometric variables such as stature, cephalic index and nasal index etc. and on certain genetic marker were available in literature on a number of Austro-Asiatic tribes, no comprehensive analysis has yet been attempted. Therefore, in order to empirically ascertain the nature and extent of genetic heterogeneity among the Austro-Asiatic groups, we tried to collate published data on these groups and have attempted a comprehensive analysis of the anthropometric and traditional genetic marker data. We have also included few transitional groups, which were once considered to be Austro- Asiatic and at present are identified with the neighboring non Austro-Asiatic groups. For the sake of comparison we have further included few non Austro-Asiatic populations, which have geographic proximity to the Austro-Asiatic tribes.

Table 1. Geographical distribution of Austro- Asiatic tribes in India.

Sl. No.	Name of population	Distribution
1	Agharias (T)	Uttar Pradesh;Madhya Pradesh, Mandla, Bilaspur, Rewa, Maikal hills; Bihar
2	Bado Gadaba (T)	Andhra Pradesh, Visakhapatnam District; Orissa, Koraput
3	Bhumij (T)	Bihar, Ranchi Jagdalpur; Orissa, Koraput
4	Dudh Kharia (T)	Jharkhand, Ranchi; Madhya Pradesh, Raigarh, Jashpur; Orissa, Sundargarh
5	Ho (T)	Jharkhand Singhbhum; Orissa, Mayurbhanj and Koenjhar; West Bengal
6	Juang (T)	Orissa, Southern Keonjhar, Northern Angul, and

		Eastern Dhenkanal
7	Khasi (T)	Assam; Meghalaya, Khasi-Jaintia hills
8	Korku (T)	Maharashtra, Amravati; Madhya Pradesh
9	Lanjia Saora (T)	Orissa, Ganjam Koraput and Phulbani; Andhra Pradesh, Srikakulam
10	Lodha (T)	Orissa, Mayurghanj and Baleswar; West Bengal, Midnapore
11	Munda (T)	Jharkhand, Ranchi; Orissa Mayurbhanj
12	Nicobarese (T)	Nicobar Island
13	Pareng Gadaba (T)	Orissa, Koraput
14	Santhals (T)	Bihar, Santhal Pargana; Orissa; West Bengal

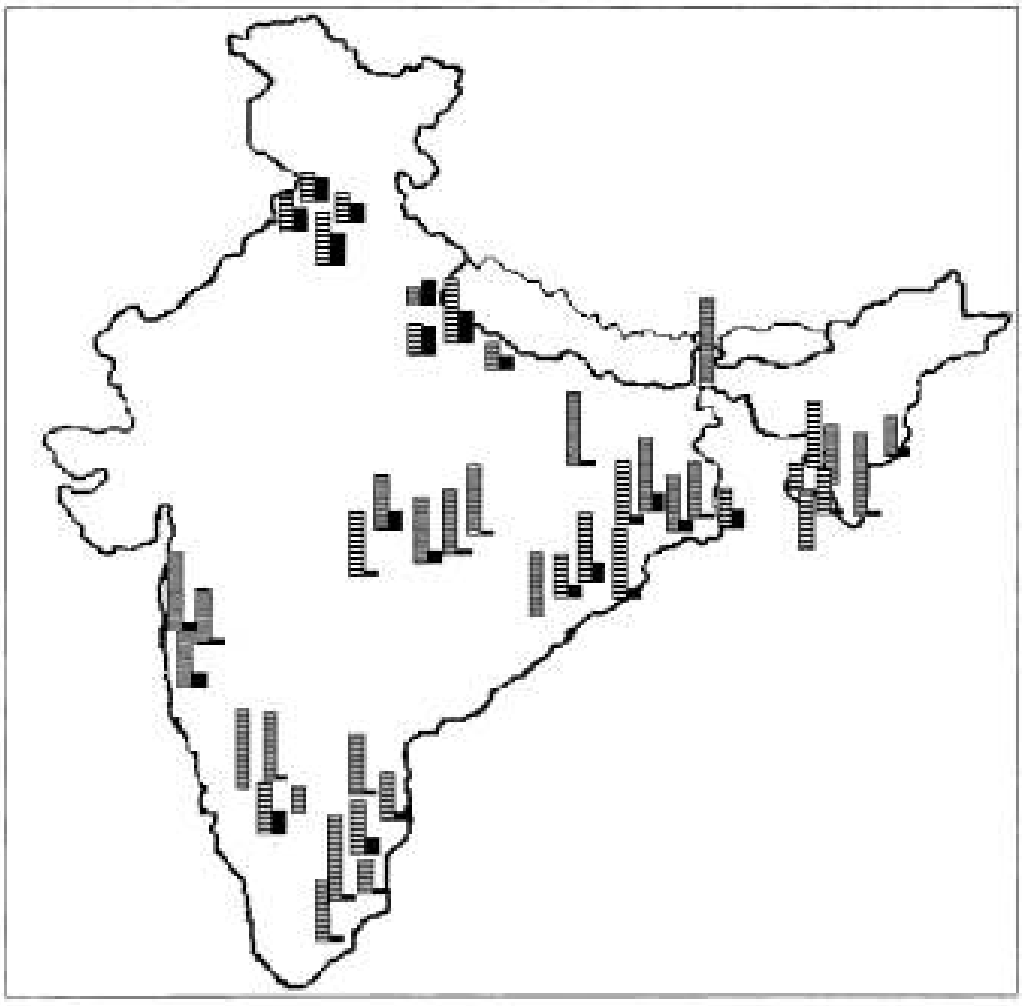
.

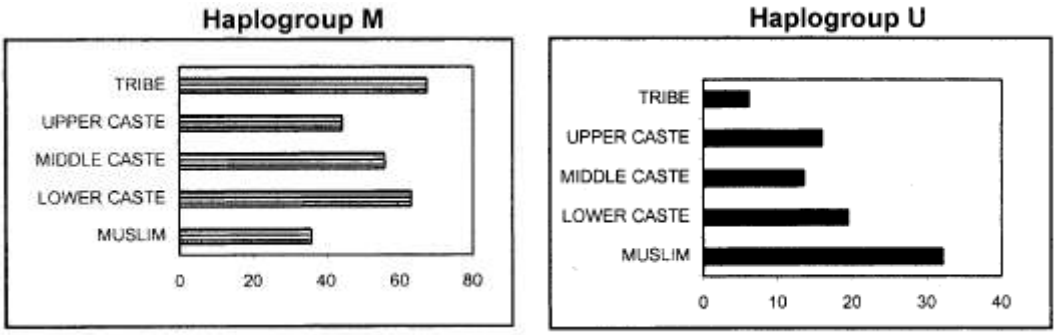
Tribes are the oldest inhabitant of India

The tribals are possibly the original inhabitants of India (Thapar 1966; Ray 1973), although their evolutionary histories and biological contributions to the nontribal populations have been debated (Risley 1915; Guha 1935; Sarkar 1958). Therefore, it is crucial to carry out genetic investigations in geographically and culturally disparate, but ethnically well-defined, populations, using data on a uniform set of mitochondrial (mt), Ychromosomal, and autosomal DNA markers. Unfortunately, the vast majority of earlier studies on Indian populations have been conducted on ethnically ill-defined populations or have been restricted to a single geographical area or a single set of markers— primarily either mitochondrial or Y-chromosomal (e.g.,Kivisild et al. 1999a; Bamshad et al. 2001). The objectives of the present study are to (1) provide a comprehensive view of genomic diversity and differentiation in India, and (2) to draw inferences on the peopling of India, and the origins of the ethnic populations, specifically in relation to the various competing hypotheses, such as whether the Austro-Asiatic or the Dravidian-speaking tribal groups were the original inhabitants of India (Risley 1915; Guha 1935; Sarkar 1958).

Haplogroup Distributions

The frequencies of the most predominant mtDNA HGs in India, M and U,are roughly inversely correlated (Fig. 1). HG-M frequency is very high (overall 59.9%: range 18.5% [Brahmins of Uttar Pradesh] to 96.7% [Kota]), confirming that it is an ancient marker in India. HG-M frequency is the highest among tribal groups, particularly in the AA tribals. Among HG-M individuals, 98.22% belong to subHG-M*, defined by the presence of T at np 16,223. Figure 2A presents the frequencies of various known (Bamshad et al. 2001) subHGs of M* among different sociocultural categories. (Detailed data are given in Suppl. Tables 2–4.) Individuals belonging to subHG-M2 had the highest nucleotide diversity in HVS1, indicating that M2 may be the most ancient in India. It occurs in significantly higher ($p < 0.05$) frequencies among tribals (28%), particularly among the AA tribals (32%), than among castes (8.8%).

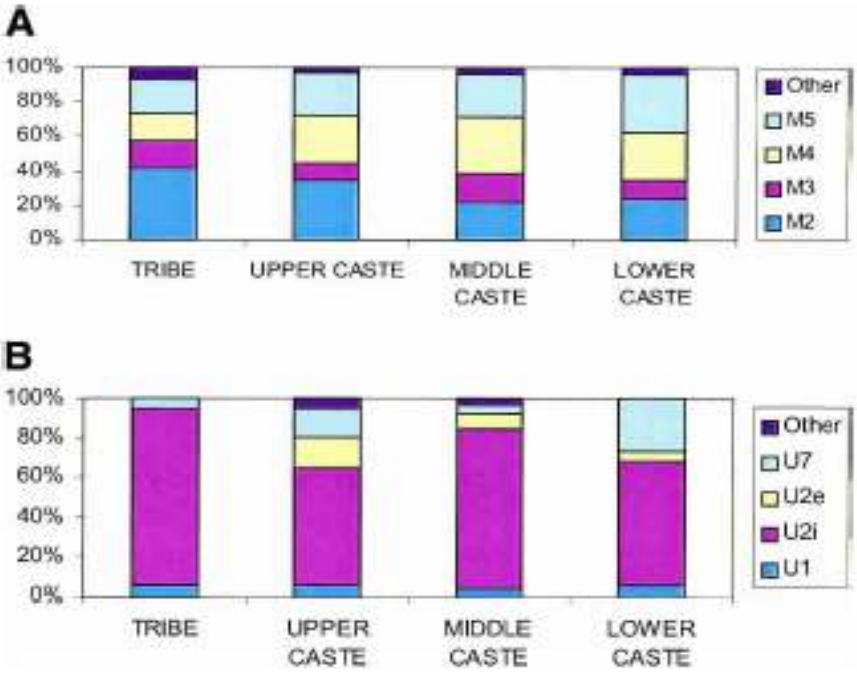




Adapted from Basu et al (2003)

Figure 4 Frequencies (%) of mitochondrial haplogroups M (hatched) and U (solid black) in 44 ethnic populations, and among sociocultural groups of populations (*insets*).

Furthermore, the coalescent time of M2 found in India was estimated to be greater than most east Asian and Papuan branches of HG-M (Forster et al. 2001), indicating that India was settled early after humankind came out of Africa (Kivisild et al. 1999b). These findings imply that the contemporary tribals are descendants of the initial settlers. HG-U is a complex mtDNA lineage, whose age was estimated from (Basu et al) data to be 45,000 _ 25,000 years, not significantly different from an earlier estimate (Torroni et al. 1996).



Adapted from Basu et al (2003)

Fig. 5. Frequencies (%) of subhaplogroups of (*A*) M and (*B*) U among tribal and ranked caste populations.

Its frequency is significantly ($p < 0.001$) higher among the IE-speaking caste groups, compared with other caste or tribal groups. Of particular interest are the frequencies of subHGs U2i (Indian-specific cluster of subHG U2 that predated the arrival of IE speakers from Central and West Asia into India; Kivisild et al. 1999a), U2e (Western-Eurasian cluster of U2),and U7 (an ancient Indian subHG). The frequencies of the subHGs of U are presented in Figure 4B. It is striking that the tribals do not possess U2e,and have the highest frequency of U2i. The radiations in frequencies (Figs. 4 and 5) of HG-M, particularly of subHG-M2, and also of HG-U, notably the absence of U2e among tribals, indicate that (1) tribals are more ancient than the castes,(2) there has been considerable admixture with Central and West Asians during the formation of the caste system, and (3) many new female lineages were introduced by the IE speakers.

Austro Asiatic tribes are the original inhabitant of India

Sociocultural and linguistic evidence indicates (Risley 1915; Thapar 1966; Pattanayak 1998) that the AA tribals are the original inhabitants of India. Some other scholars have, however, argued that tribal groups speaking DR and AA languages have evolved from an older original substrate of proto-Australoids (Keith 1936), whereas the TB tribals are later immigrants from Tibet and Myanmar (Guha 1935). Basu et al findings strongly support the hypothesis that AA tribals are the earliest inhabitants of India. They possess the highest frequencies of the ancient east- Asian mtDNA HG-M and exhibit the highest HVS1 nucleotide diversity (Table 2). They also have the highest frequency of subHG M2 (19%),which had the highest HVS1 nucleotide diversity compared with other subHGs and therefore possibly the earliest settlers (the estimated coalescence time is 63,000 _ 6000 ybp; Kivisild et al. 1999a). Although all sociolinguistic groups seem to have undergone significant population expansions as evidenced by the unimodality of the HVS1 mismatch distributions (data not shown) and by the values of the relevant statistics (small values of the "raggedness" statistic and significantly large negative values of Fu's *F*s Statistic; Table 2),the AA tribals show the highest value of the estimated expansion time, 55,000 years, which is <15,000 years larger than the estimates for the other groups. Although we cannot be sure that this expansion took place in India, in conjunction with the other findings, it appears that this group of tribals may be the earliest inhabitants of India. A young subclade M4,with an estimated coalescence time of 32,000 _ 7500 ybp (Kivisild et al. 1999a),whose overall frequency is <15% in India, is completely absent among them. It is, therefore, likely that M4 arose after the expansion of the AA tribals and their entry into India.

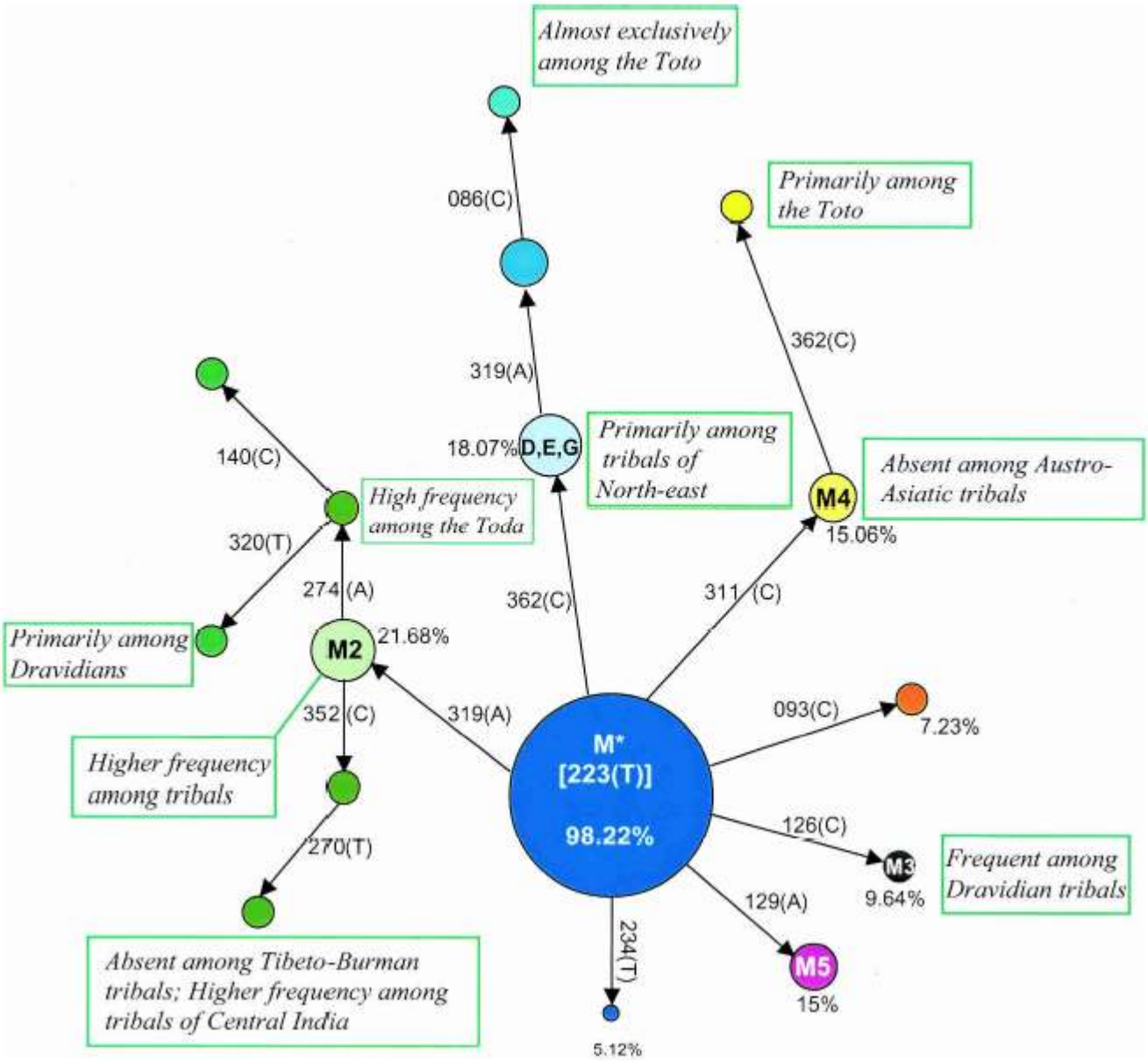


Fig. 6. Phylogenetic network of mtHVS1 sequences belonging to subhaplogroup M*, with frequency distributions of motifs in populations.

Table 2. Number of Polymorphic Site, Nucleotide Diversity, Mismatch Statistics and Estimated Expansion Times, other statistics to Population Expansion Based on HVS1 sequence data

	Linguistic group		
	Austro Asiatic	Dravidian	Tibeto-Burman
No. of sequences	46	94	95
No. of polymorphic site	57	55	74
Nucleotide Diversity	0.0224	0.0170	0.0173
Mean No. of mismatches	8.00	6.07	6.16
Expansion time (y)	54,656	41,470	42,085
Fu's Fs (p value)	-24.93 (0.0)	-25.26	-25.20

Y-chromosomal data also support that the Austro Asiatic speaking people came first in India. High frequencies of Y-HG K* (Fig. 3) are found among the TB populations, mainly confined to northeast India, and also among the Han Chinese (Su et al. 2000). The TB subfamily of the Sino-Tibetan language family has been subdivided (Grimes 1999) into four branches: Baric, Bodic,Burmese-Lolo,and Karen. Based on a study of Y-chromosomal haplotypes,Su et al. (2000) have contended that after the proto-Tibeto-Burman people left their homeland in the Yellow River basin, the Baric branch moved southward and peopled the northeastern Indian region after crossing the Himalayas. This branch did not possess the YAP insertion element, which has also not been found in any of the TB populations of India. TB speakers entered India from the northeastern corridor.

Table 3. Frequencies of Y Haplogroups in Ethnic Populations of India Belonging to Various Sociolinguistic Groups

Linguistic group	**Social group**	**Sample size**	**Haplogroup (%)**					
			P*	**BR***	**R1a**	**J**	**K***	**L**
Austro- Asiatic	Tribe	52	7.7	28.8	0.0	13.5	50.0	0.0
Dravidian	Tribe	84	4.8	67.9	3.6	9.6	6.0	8.2
	Caste	103	19.4	32.2	11.6	22.3	4.8	9.6
Tibeto Burman	Tribe	87	4.7	10.3	0.0	14.9	70.1	0.0
Indo-European	Tribe	19	0.0	36.8	15.8	21.1	26.3	0.0
	Caste	122	22.1	32.8	23.8	12.3	6.6	2.4

Austro – Asiatic tribes, who primarily inhabit the eastern and central Indian regions, also possess high frequencies of Y-HG K* (Fig. 7). There is one major tribal group (Khasi) of northeastern India, who speak a dialect that belongs to the AA subfamily. Besides India, AA languages are spoken in south-east Asia. Thus, it is likely that a fraction of the AA tribals also entered India through the northeastern corridor. However, it does not seem that all of them have entered through this corridor. The expansion time of AA tribals was estimated to be <55,000 yr (Table 4), which is <13,000 yr greater than that estimated for TB tribals. But the age of the Y-HG K* estimated from the pooled variance of repeat numbers at the STRP loci among AA tribals (8500 yr) is about half of that estimated for TB tribals (15,000 yr).

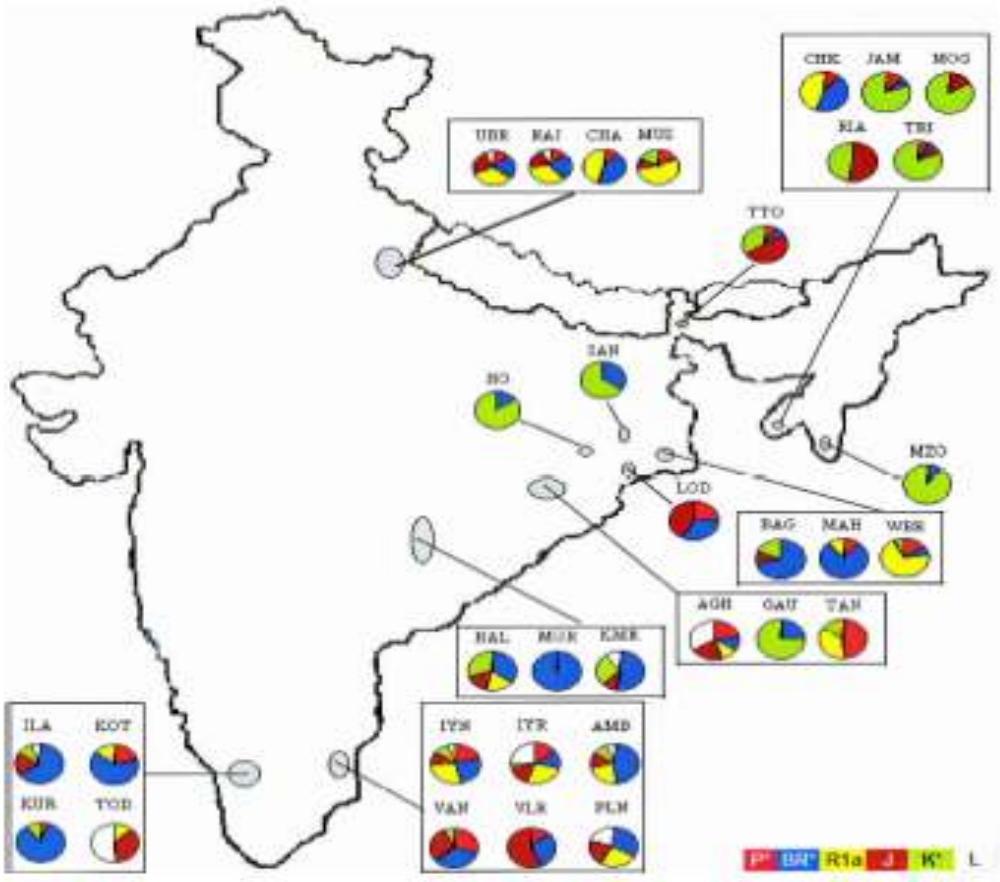


Fig.7. Frequencies (%) of Y-chromosomal haplogroups among ethnic populations. (Population Name and codes are given bellow.)

1. Agharia [AGH] 2. Ambalakarer [AMB] 3. Bagdi [BAG] 4. Chakma [CHK] 5. Chamar [CHA] 6. Gaud [GAU] 7. Gond [GND] 8. Halba [HAL] 9. Ho [HO] 10. Irula [ILA] 11. Iyengar [IYN] 12. Iyer [IYR] 13. Jamatiya [JAM] 14. Jat Sikh [JSK] 15. Kamar [KMR] 16. Khatris [KHT]17. Konkan Brahmins [KBR] 18. Kota [KOT] 19. Kurumba 20. Lodha [LOD] 21. Mahishya [MAH]22. Manipuri (Meitei) 23. Maratha [MRT] 24. Mizo [MZO] 25. Mog [MOG] 26. Munda [MUN]27. Muria [MUR] 28. Muslim [MUS] 29. Naba-Baudh [NBH] adopted Buddhism) 30. Pallan [PLN]31. Punjab Brahmins [PBR] 32. Rajput [RAJ] 33. Riang [RIA] 34. Santal [SAN] 35. Saryupari Brahmins [SBR] 36. Scheduled caste-Punjab [SCH] 37. Tanti [TAN] 38. Tripperah (Tripuri) [TRI] 39. Toda [TOD] 40. Toto [TTO] 41. Uttar Pradesh Brahmins [UBR] 42. Vanniyar [VAN] 43. Vellala [VLR] 44. West Bengal Brahmins [WBR]

This is possible that the ancestors of the AA speakers entered India through the northwest from out-of-Africa as they moved south of the Himalayas, and another ancestral group moved north of the Himalayas, settled in southern China, and then entered later through the northeast.

Conclusions

The migration routes of modern humans into India remain enigmatic, and whether there were also returns to Africa from India/Asia is unclear (Maca-Meyer et al. 2001; Roychoudhury et al. 2001; Cruciani et al. 2002). Some scholar said that Austro-Asiatic languages speaking tribes probably came from Africa, entered India through western corridor (Nei and Ota 1991; Chu *et al* 1998; Su *et al* 1999; Majumder 2001). There is an alternative hypothesis in vogue, which indicates that these groups have migrated from China into India through northeastern corridor (Ballinger *et al* 1992; Gadgil *et al* 1998; Diamond 1997). But the genetic data indicate that the Austro Asiatic tribes are the initial settler of India , it is exhibited higher genetic diversity in compare to the other linguistic group and simultaneously it have more M2 haplogroup which is a sub- haplogroup of M.

The Austro Asiatic tribes exhibited highest value of estimated expansion time, 55,000 years, which is 14,000 years larger than the estimated for the other groups, which indicate that this group of tribal may be the earliest inhabitant of India.

Y chromosomal data also indicate that, Austro Asiatic tribes may be initial settler of India Because is shows more K* haplogroup in compare to other logistic group except Tibeto-Burman Linguistic family. These data indicate that the Austro Asiatic tribes may come from East Asia. Our data also support this, Munda and Birhor are the two tribal population of Jharkhand which belongs to the Austro Asiatic linguistic family which have high frequency of "O "haplogroup caring M-132.

References

1. Cann, R.L. 2001. Genetic clues to dispersal of human populations: Retracing the past from the present. *Science* **291:** 1742–1748.
2. Maca-Meyer, N., Gonzalez, A.M., Larruga, J.M., Flores, C., and Cabrera, V.M. 2001. Major genomic mitochondrial lineages delineate early human expansions. *BMC Genetics* **2:** 13–0.
3. Singh, K.S. 1992. *People of India: An introduction.* Anthropological Survey of India, Calcutta, India.
4. Kosambi,D.D. 1991. The culture and civilization of ancient India in historical outline. Vikas Publishing House,New Delhi,India.
5. Ray, N. 1973. *Nationalism in India.* Aligarh Muslim University,Aligarh, India.
6. Thapar,R. 1966. *A history of India,* Volume 1. Penguin,Middlesex,UK. 1995. The first millennium B.C. in northern India (up to the end of Mauryan period). In *Recent perspectives of early Indian history* (ed. R. Thapar),pp. 80–141. Popular Prakashan,Bombay,India. 2003. *Early India: From the origins to AD 1300.* University of California Press,Berkeley,CA.
7. Risley,H.H. 1915. *The people of India.* Thacker Spink,Calcutta,India.
8. Guha,B.S. 1935. The racial affinities of the people of India. In *Census of India, 1931*,Part III Ethnographical. Government of India Press, Simla,India.
9. Sarkar,S.S. 1958. Race and race movements in India. In *The cultural heritage of India* (ed. S.K. Chatterjee),Vol. 1,pp. 17–32. The Ramakrishna Mission Institute of Culture,Calcutta,India.
10. Buxton L H D 1925 *The peoples of Asia* (London) Cann R L 2001 Genetic clues to dispersal of human populations: Retracing the past from the present; *Science* **291** 1742–1748
11. Pattanayak D P 1998 The language heritage of India; in *The Indian human heritage* (eds) D Balasubramanian and N A Rao (Hyderabad: Universities Press) pp 95–99

12. Renfrew C 1987 *Archaeology and language: The puzzle of Indo-European origins* (London: Jonathan Cape) Renfrew C 1992 Archaeology, genetics and linguistic diversity; *Man* **27** 445–487

13. Ruhlen M 1991 *A guide to the world's languages* (Stanford: Stanford University Press)

14. Majumdar P P 2001 Ethnic populations of India as seen from an evolutionary perspective; *J. Biosci*. **26** 533–545

15. Majumdar P P, Roy B, Banerjee S, Chakraborty M, Dey B, Mukherjee N, Roy M, Thakurta P G and Sil S K 1999 Human-specific insertion/deletion polymorphisms in Indian populations and their possible evolutionary implications; *Eur*. *J*. *Hum*. *Genet*. **7** 435–446.

16. Ballinger S W, Schurr T G, Torroni A, Gan Y Y, Hodge J A, Hassan K, Chen K H and Wallace D C 1992 Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations; *Genetics* **130** 139–152.

17. Lal B B 1956 Paleolithics from the Beas and Banganga valleys Punjab; *Ancient India* **12** 58–92.

18. Mohapatra G C 1975 Acheulian Element in Soan Culture Area; *J*. *Archaeol*. *Soc*. *Nippon* **60** 4–18

19. Gupta S P 1979 Archaeology of Soviet Central Asia and the Indian borderlands (Delhi: B R Publishing)

20. Nei M and Ota T 1991 Evolutionary relationships of human population at molecular level; in *Evolution of life fossils molecules and culture* (eds) S Osawa and T Honjo (Heidelberg: Springer-verlag) pp 415–428.

21. Chu J Y, Huang W, Kuang S Q, Wang J M, Xu J J, Chu Z T, Yang Z Q, Lin K Q, Li P, Wu M, Geng Z C, Tan C C, Du R F and Jin L 1998 Genetic relationship of populations in China; *Proc*. *Natl*. *Acad*. *Sci*. *USA* **95** 11763–11768.

22. Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Luo J, Chu J, Tan J, Shen P, Davis R, Cavalli-Sforza L L, Chakraborty R, Xiong M, Du R, Oefner P, Chen Z and Jin L 1999 Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age; *Am*. *J*. *Hum*. *Genet*. **65** 1718–1724

23. Vikrant Kumar and B Mohan Reddy 2003 Status of Austro Asiatic groups in the peopling of India:

24. Cavalli –Sfroza LL,Menozzi P and Piazza A 1994 The history and geography of human genes(Princeton,NJ: University Press)

25. Ray,N. 1973. *Nationalism in India.* Aligarh Muslim University, Aligarh, India.

26. Bamshad,M.J.,Kivisild,T.,Watkins,W.S.,Dixon,M.P.,Ricker,L.E.,Rao,B.B.,Naidu, M.,Prasad,B.V.R.,Reddy,P.G.,Rasanayagam,A., et al. 2001. Genetic evidence on the origins of Indian caste populations. *Genome Res.* **11:** 994–1004.

27. Kivisild,T.,Bamshad,M.J.,Kaldma,K.,Metspalu,M.,Metspalu,E.,Reidla,M.,Laos,S., Parik,J.,Watkins,W.S.,Dixon,M.E.,et al. 1999a. Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr. Biol.* **9:** 1331–1334.

28. Forster, P., Torroni, A., Renfrew, C., and Rohl,A. 2001. Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Mol. Biol. Evol.* **18:** 1864–1881.

29. Torroni,A.,Huoponen,K.,Francalacci,P.,Petrozzi,M.,Morelli,L.,Scozzari,R.,Obinu, D.,Savontaus,M.-L.,and Wallace,D.C. 1996.Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144:** 1835–1850.

30. Risley,H.H. 1915. *The people of India.* Thacker Spink,Calcutta,India.

31. Grimes,B.F. 1999. *The ethnologue: Languages of the world.* Summer Institute of Linguistics,California.

32. Su,B.,Xiao,C.,Deka,R.,Seielstad,M.T.,Kangwanpong,D.,Xiao,J.,Lu,.,Underhill,P.,C avalli-Sforza,L.L.,Chakraborty,R.,et al. 2000. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas. *Hum. Genet.* **107:** 582–590.

33. Analabha Basu,1,4 Namita Mukherjee,1,4 Sangita Roy,2,4 Sanghamitra Sengupta,1,4 Sanat Banerjee,1 Madan Chakraborty,1 Badal Dey,1 Monami Roy,1 Bidyut Roy,1 Nitai P. Bhattacharyya,3 Susanta Roychoudhury,2 and Partha P. Majumder1,5 2003 , Ethnic India: A Genomic View, With Special Reference to Peopling and Structure, Genome Research 13:2277–2290

34. Partha P Majumdar 2001, Ethnic population of India as seen from an evolutionary perspective, J. Bioscience 26: 533-545.

35. Thangaraj K, Sridhar V, Kivisild T, Reddy AG, Chaubey G, Singh VK, Kaur S, Agarawal P, Rai A, Gupta J, Mallick CB, Kumar N, Velavan TP, Suganthan R, Udaykumar D, Kumar R, Mishra R, Khan A, Annapurna C, Singh L 2005. Different population histories of the Mundari- and Mon-Khmer-speaking Austro-Asiatic tribes inferred from the mtDNA 9- bp deletion/insertion polymorphism in Indian populations. *Hum Genet*, **116:** 507-17.

36. Reddy BM, Langstieh BT, Kumar V, Nagaraja T, Reddy ANS, et al (2007) Austro-Asiatic Tribes of Northeast India Provide Hitherto Missing Genetic Link between South and Southeast Asia. PLoS ONE 2(11): e1141. doi:10.1371/journal.pone.0001141.

37. Campbell MC, Tishkoff SA (2008), African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping .Ann Rev Genomics Hum Genet 9: 403-33.

38. Morelli L, Contu D, Santoni F, Whalen MB, Francalacci P, Cucca F(2010) A comparison of Y-chromosome variation in Sardinia and Anatolia is more consistent with cultural rather than demic diffusion of agriculture. PLoS One 29;5(4):e10419