

SOFTWARE REVIEW

Of fine pedigree

Cyrillic, version 2.1. Platform: PC, Windows 3.1, 4 Mb RAM [Mac: Designed for Windows but can be used on PowerMac running SoftWindows (see <http://www.insignia.com/SoftWindows>)]

Cherwell Scientific, Magdalen Centre, Oxford Science Park, Oxford OX4 4GA, UK (<http://www.cherwell.com>).
India (authorized distributor): KEMS Software, 725 10th Main, 4th Block, Jayanagar, Bangalore 560 011 (e-mail: kems@bgl.vsnl.net.in). Price in India Rs 44,000.

Reviewed by N. V. JOSHI*

A typical example of the legendary British (at least of yesteryears) penchant for understating accomplishments, Cyrillic goes far beyond mere pedigree drawing. It is a very powerful tool for organizing and managing complex genetic information for a large family tree, in a convenient, user-friendly manner.

Study of genetic disorders (which naturally involves pedigree analysis) is one of the most active and important areas of contemporary research in genetics today. An additional impetus to such investigations has been provided by the discoveries revealing genetic predisposition as an important component of susceptibility of individuals to non-genetic diseases as well. However, most such links between disease susceptibility and genetic background are subtle and weak, and are often rather difficult to detect owing to the large number of masking and confounding factors. Data from a large number of genetically related individuals are essential for such studies. The recent advances in molecular biology have revolutionized the techniques for rapidly screening a large number of individuals for presence of specific genetic markers. These relatively inexpensive methods generate large volumes of data, which then need to be subjected to long and tedious analysis. In many situations, a brute-force computation is neither possible nor appropriate for detecting patterns in such complex data sets, and a careful scrutiny by experienced practitioners is necessary to 'make sense' of the information.

Drawing pedigrees is one of the best methods (and often the only one) for presenting such genetic information. This is not very difficult if all you have to handle is a family tree spanning a few generations, a few genotypes/traits, and maybe several dozen individuals. However, even in such a relatively simple situation, modifying or updating information about the tree is at best cumbersome and time-consuming;

a rapid display of several versions of a tree is simply not possible. This is exactly where Cyrillic steps in. For starters, it can handle up to 10,000 individuals, and up to 250 markers per chromosome at a time, and can show you up to nine family trees simultaneously. Add to it the cut-and-paste editing facility for transferring information between family trees, provision to present details at various levels, and excellent colour-coding and printing options, and you get a glimpse of the power of Cyrillic.

Traditionally, use of computers and software has been mainly for numerical computations—either involved calculations, or analysis of large data sets. The last one and a half decade has brought graphics and database management to the fore. The current buzzword is visualization, and that is what Cyrillic is about. In fact, one can imagine it being advertised as 'the BEST multidimensional visualization software' in screaming, multicolour headlines had it been developed across the Atlantic.

Now to the details of the product itself, though it is a little odd to review in 1998 a product developed for Windows 3.1, and distributed on a single 1.44-Mb floppy diskette (and a non-write-protected one at that)! However, in terms of ease of use and power, Cyrillic (version 2.1) is as good as (if not better than) anything I have come across. Installation is extremely straightforward, and over in a matter of minutes. It takes only about 4Mb of disk space (spread over more than 45 different directories, however), and loads very fast (on a 200-MHz Pentium with 32 Mb RAM). The software comes with a neatly spiral-bound, beautifully printed manual, which is very user friendly, and by scanning the first few pages one is able to begin using the software in virtually no time. The accompanying quick-reference card is also a thoughtful and useful addition.

Bundled along with the software are some examples (including one of European royalty, with the curious phrase 'intellectual disability' thrown in as a comment), very useful online help, and, most importantly, four tutorial files. With

*Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560 012, India. E-mail: nvjoshi@ces.iisc.ernet.in.

their help, within half an hour or less, anyone would be confident enough of undertaking the task of drawing even the most complex genealogies. The procedures for creating new family trees and then modifying or editing them are very intuitive. Almost anything can be done just with a few clicks of the mouse (yet another instance where mouse-based systems have been of great help to geneticists!)—adding spouses/children, adding biographical information, adding genetic information, and so on. A table-based data entry feature makes matters even simpler.

The real power of Cyrillic, however, is in the way it handles genetic markers. Traits, diseases, haplotypes, etc. can be very easily added, and, more importantly, can be very distinctively represented using different colours and styles. A perpetual problem in drawing family trees is the choice of the level of detail. On the one hand, one would like to prepare as comprehensive a data set as possible—this is essential for detecting some not-so-obvious connections. On the other hand, the drawing tends to become more and more complex, to the extent of being unintelligible as more traits are added. Cyrillic neatly finesses the problem by providing unlimited flexibility in the choice of detail that you wish to represent at any given time. A click of the mouse gives a menu, and you can mix and match at will. The tree is instantly drawn to enable you to explore a large number of combinations till you zoom in on the best one (and yes, you can zoom in and zoom out to change the size as well).

Cyrillic has one of the most flexible sets of printing options that I have seen. It allows you to customize the drawing at the preview stage in various ways, very easily; particularly noteworthy is the facility to change the size of the drawing in steps of 5%. Both colour and black-and-white drawings come out very well—ready to be sent for publication. There was, however, a slight mismatch between the 'preview' and what actually came out in the printout. On the positive side, a cluttered-looking and overlapping box (in the preview) describing the meanings of symbols was well separated in the printout. On the negative side, the landscape versus portrait orientations did not match between preview and printout, and the drawing was unnecessarily spread over more than one sheet (this also showed the capability of the software to print very large trees by

spreading them over several sheets). There was also a slight difference between the selected size and the obtained one. However, with a little more familiarity with the software, these minor irritants can easily be avoided.

Finally, it is not just data compilation, editing, visualization and printing—Cyrillic can carry out some useful analysis as well. It can automatically detect consanguinity; it can calculate inbreeding and kinship coefficients; and it can perform risk assessment as well. More importantly, it allows export of data in a form that can be readily accepted by other, more sophisticated and special-purpose programs for analysis. It can also import files from several other formats. The manual gives a very detailed description of the formats of the files used by Cyrillic. This is just another instance where one can see that development of this software is more a labour of love, something created by one experienced scientist for others of the clan, and not something commissioned after a careful market survey and with an eye on the balance sheet.

Support for the software is advertised on the website of the publishers (www.cherwell.com), and that is where one can obtain pricing and ordering information as well. In step with the shift towards e-commerce, you can order it on the web itself—a form-based interface allows you to key in the details (nearest dealer, whether government/academic/commercial buyer, etc.) and generates a price statement for you. For both academic and government users from the 'Rest of the world', Cyrillic costs 449 pounds—about Rs. 31,000 as of December 1998 [But note the Indian price provided by the authorized Indian distributor: Rs. 44,000.—Ed.]. It is difficult to decide whether this means 'affordable' or 'too high'. On the one hand, the 'open software' culture has released far more powerful, complex and useful packages (the Linux operating system, the GNU C and C++ compilers, for example) absolutely free of charge. On the other hand, especially in clinical situations, even if one patient is helped by use of this software (which can very well happen), the price is trivial in comparison with the return. On the whole, for laboratories and departments where pedigree drawing and analysis are needed on a routine basis, it would be very worthwhile to invest in Cyrillic 2.1, as matters stand today.

THE INDIAN HUMAN HERITAGE

Editors

D Balasubramanian

N Appaji Rao



Universities Press

12

Peopling of India

Madhav Gadgil, N.V. Joshi
Centre for Ecological Sciences
Indian Institute of Science, Bangalore

U.V. Shambu Prasad
Centre for Research in Indo-Bangladesh Relations
Jodhpur Park, Calcutta

S. Manoharan, Suresh Patil
Anthropological Survey of India
Southern Regional Centre, Mysore

Abstract

We examine the demographic history of India on the basis of a new investigation of mitochondrial DNA base sequences of 101 Indians, in light of the recent synthesis of global genetic history of humans by Cavalli-Sforza and co-workers. To this population genetic evidence we add fresh insights into linguistic and anthropological patterns based on the People of India Project of the Anthropological Survey of India, and a review of the pertinent archaeological evidence on waves of diffusion of technological innovations over the subcontinent. The Indian subcontinent has been populated by a series of migrations propelled by significant technological innovations outside India since the first major expansion of non-African *Homo sapiens*, probably around 65,000 years before present (BP). The likely major migrations include (i) Austric language speakers soon after 65,000 years BP, probably from the northeast; (ii) Dravidian speakers around 6,000 years BP from the mideast with knowledge of cultivation of crops like wheat, and domestication of animals like cattle, sheep and goats; (iii) Indo-European speakers in several waves after 4000 years BP with control over horses and knowledge of iron technology; and (iv) Sino-Tibetan speakers in several waves after 6000 years BP with knowledge of rice cultivation. A notable feature of Indian society is the persistence of thousands of tribe-like endogamous groups in a complex agrarian and now industrial society. In this society, populations of dominant groups have continued to grow, while those of subjugated groups may have stagnated most of the time.

Introduction

India is a country remarkable for its diversity, both biological and human. The biological diversity owes itself to the country's position at the trijunction of the African, the northern Eurasian and the Oriental realm, its great variety of environmental regimes, and its relative stability of biological production. It is this biological wealth that has attracted to the subcontinent many streams of people at different times, from different directions, bringing together a great diversity of human genes and human cultures. While in other lands the dominant human cultures have tended to absorb or eliminate others, the tendency in India has been to isolate and subjugate the subordinated cultures, thereby

augmenting cultural diversity. This tendency to nurture diversity has been favoured by the diversity of the country's ecological regimes [Gadgil and Guha, 1992].

People migrate because of pulls from their destination and pushes in their homeland, often propelled along by some technological advantage. Thus in the 16th century, Europeans came to India in search of spices, pushed out by the Little Ice Age that had gripped Europe, equipped with superior sea-going vessels and guns. That migration is well documented and understood; but it is the many earlier ones that have brought to India the bulk of human genes and cultural traits. It is our purpose in this paper to elucidate what we can, of these earlier migrations.

Role of innovations

People have of course migrated out of India as well, but these out-migrations have been on a much smaller scale, and mostly over the last three centuries. This is related to the fact that India has never been the site of any significant technological innovations. A series of important innovations have, over the years, taken place outside India—innovations which have given an edge to people in control of these innovations, propelling major migrations [Habib, 1992].

In chronological order, the most relevant of these include: (i) evolution of symbolic language, probably by the first modern *Homo sapiens*, in Africa, perhaps around 100 kybp (kybp = thousand years before present); (ii) husbanding of wheat, barley, cattle, pig, etc. in the mideast around 10 kybp; (iii) husbanding of rice, buffalo, etc. in China and Southeast Asia around 8 kybp; (iv) domestication of the horse in Central Asia around 6 kybp; (v) use of iron in Anatolia around 5 kybp; (vi) use of stirrups for horse riding in Central Asia around 2 kybp; (vii) use of gunpowder in China around 2 kybp; (viii) use of cannons and guns in war in Arabia in the 15th century.

Our theme then is that these manifold innovations to the west, east and north of the Indian subcontinent have propelled many waves of people into our land, giving rise to what is genetically as well as culturally the most diverse society in the world. There are diverse lines of evidence for these migrations—genetic, linguistic, archaeological, and anthropological. We will endeavour to draw on all these disciplines to reconstruct the story of the peopling of India.

Genetic affinities

Genetically and culturally, India is perhaps the most diverse country on the face of the Earth. The most authoritative summary of the genetics of human populations is provided by Cavalli-Sforza in his magnum opus, *History and Geography of Human Genes* [Cavalli-Sforza et al., 1994]. He provides global maps of frequencies of 82 genes for 42 population aggregates of indigenous people covering the entire world. The 82 loci show the highest levels of heterozygosity (0.35–0.37) for northwestern India, West Asia and continental Europe (Fig. 1). Parts of southern and eastern India share slightly lower levels of 0.33–0.35 with western China, Central Asia, Scandinavia and northern Africa. The lowest levels of 0.21–0.23 occur in New Guinea and western Australia. Such genetic data is

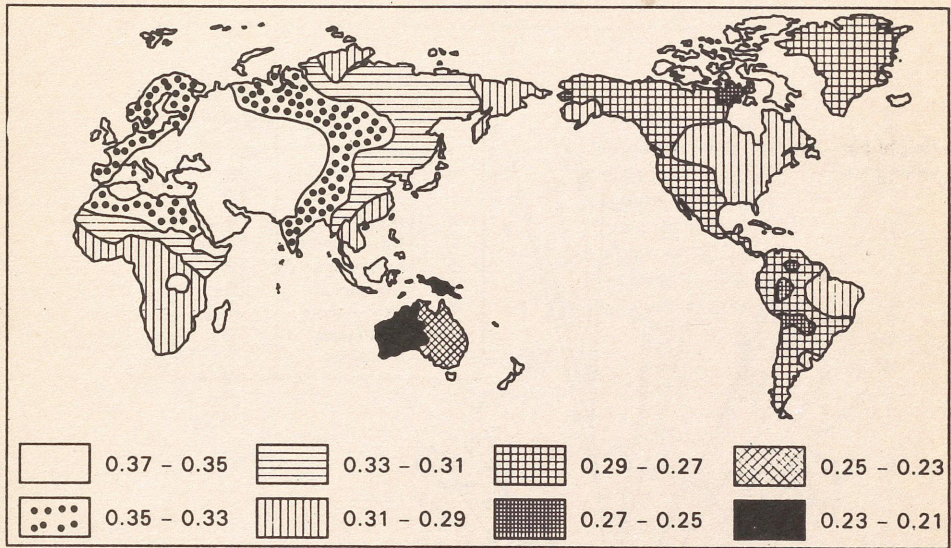


Fig. 1 Global distribution of mean genetic heterozygosity of indigenous populations based on frequencies of 82 genes for 42 world population aggregates [Cavalli-Sforza et al., 1994]

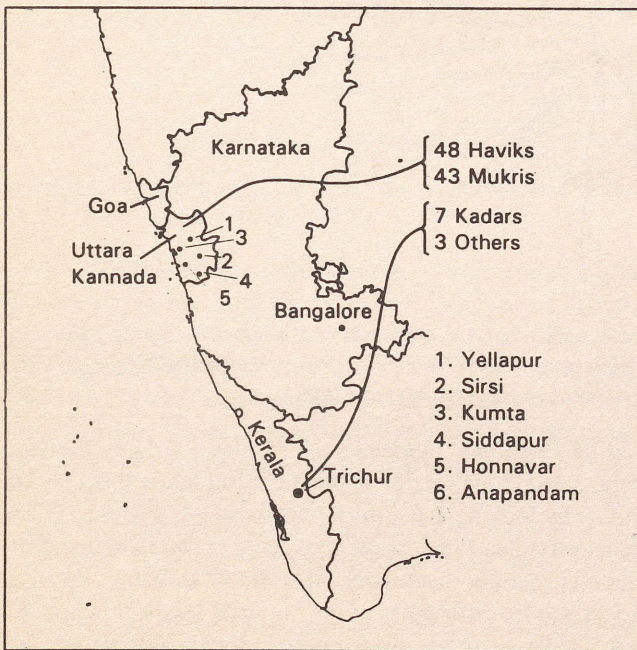


Fig. 2 Geographical location of collection of genetic material from Havik (48), Mukri (43) and 7 Kadar populations in the South Indian states of Karnataka and Kerala [Mountain et al., 1995]

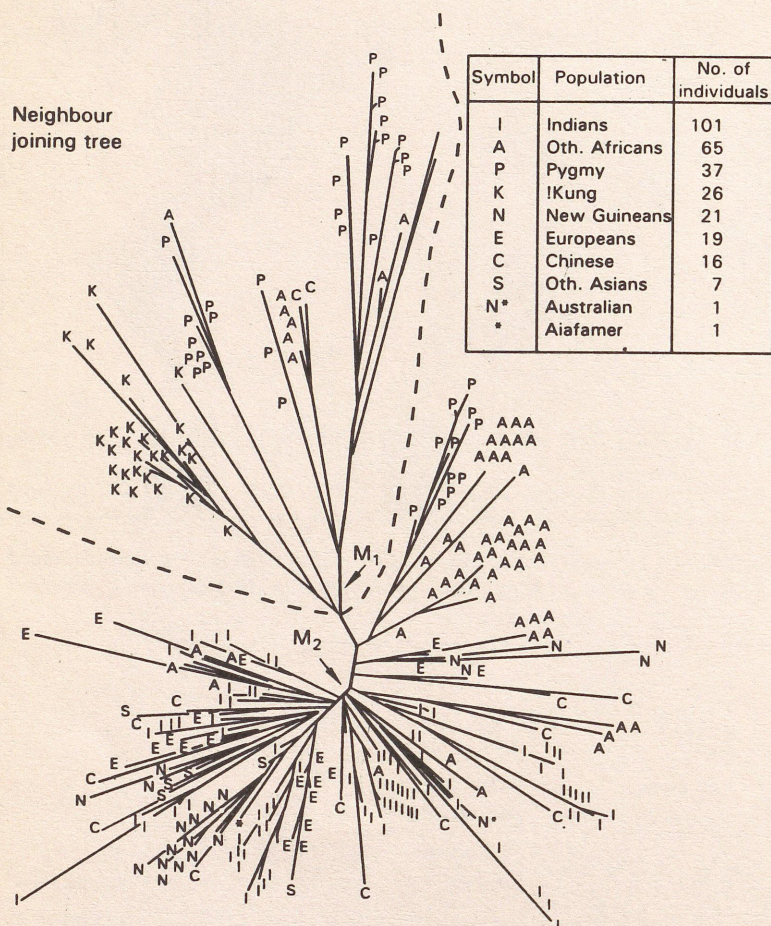


Fig. 3 Neighbour joining tree based on 745 mitochondrial DNA sequences from 101 Indians, 36 Pygmies, 26 !Kung, 65 Other Africans, 21 New Guineans, 19 Europeans, 16 Chinese, 7 Other Asians, 1 Australian and 1 Afro-American [Mountain et al., 1995]

however rather limited, based on traditional markers such as blood groups. Modern genetic techniques have greatly added to the wealth of genetic information that may be obtained from a single individual by looking at the nucleotide base sequences of DNA molecules themselves. The most variable of such sequences occur in two hypervariable regions of mitochondria, which are purely maternally inherited in humans. We have collaborated with Cavalli-Sforza and his colleagues of the Stanford Medical School to examine the DNA sequences of 791 base pair lengths from the D-loop region of mtDNA for 101 Indians [Mountain et al., 1995]. Of these, 48 belonged to the upper caste Havik group and 43 to the scheduled caste group of Mukris from the coastal Uttara Kannada

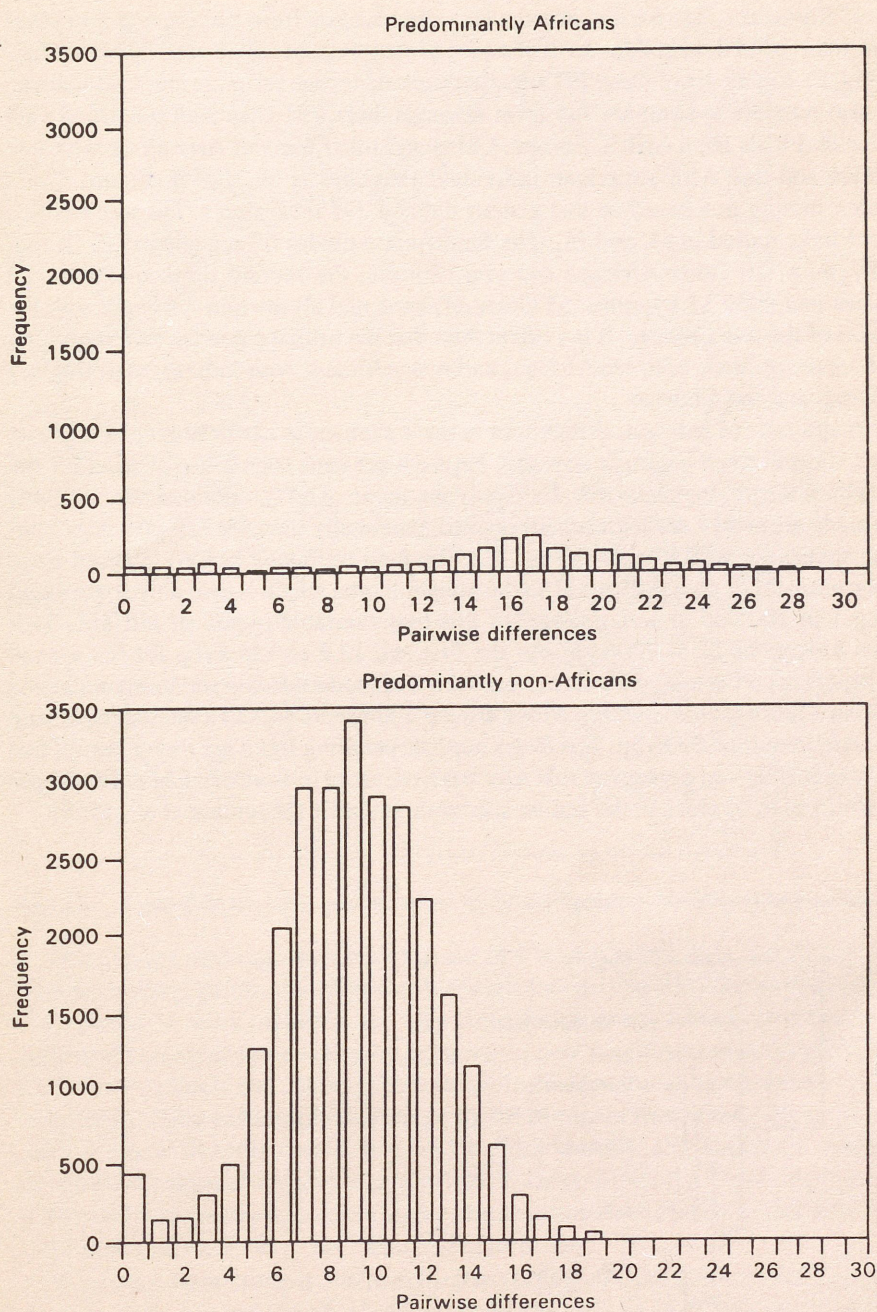


Fig. 4 Distribution of pairwise differences in the 745 mitochondrial DNA base pair sequences for the 65 sequences rooted at M_1 and 229 sequences rooted at M_2 for the neighbour joining tree depicted in Fig. 3 [Mountain et al., 1995]

region of Karnataka, 7 to a tribal population called Kadars from Kerala, and 3 to other Indians involved in field collection of samples of scrapings of cheek cells and scalp hair roots (Fig. 2). Eighty-six of these 791 sites demonstrated some variation amongst Indians; it was also possible to compare 745 from amongst these 791 sites with published data on 187 individuals from Africa, Europe, China and other parts of Asia along with one Australian and one Afro-American individual [Vigilant et al., 1991]. Figure 3 is a neighbour joining tree based on this genetic data on 294 individuals. The tree has two distinct trunks rooted in M_1 and M_2 . The first trunk includes 65 sequences: all !Kung, most Pygmies, 10 Other Africans and two Chinese; the second trunk includes 229 sequences including 11 Pygmies, 55 Other Africans and all the non-Africans with the exception of the two Chinese. It is evident then that the primary genetic differentiation of the human species is between Africans and non-Africans, with Indians intermingling with Europeans and Chinese.

The magnitude of base pair differences in these sequences can permit us to estimate the time elapsed since common ancestry. Figure 4 presents such a distribution for the two trunks of the phylogenetic tree. Evidently the group of 65 (predominantly Africans) with a mode around 17 are far more diversified genetically than the 229 primarily non-African sequences with a mode around 10. The time estimated to have elapsed since common ancestry of course depends on the mutation rate, which is probably somewhere between 1 in 100,000 to 1 in million for this hypervariable region of mtDNA. That gives us a range of 22 to 220 kybp for the first and 13.6 to 136 kybp for the second trunk. This is in conformity with the current view that modern *Homo sapiens* populations underwent a first expansion within Africa around 100 kybp, and a second expansion outside Africa around 65 kybp. The *Homo sapiens* peopling India are then a part of this second expansion—an expansion that may have occurred in southern China [Ballinger et al., 1992] or in or close to the Indian subcontinent itself [Mountain et al., 1995].

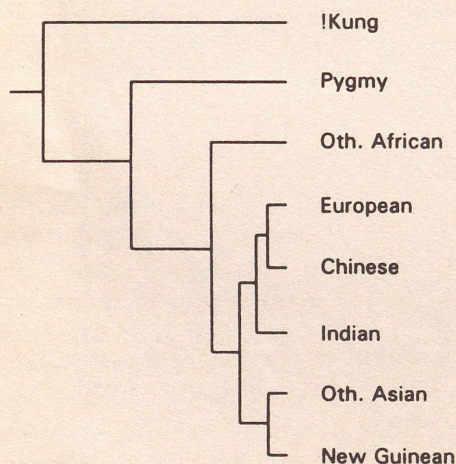


Fig. 5 Dendrogram inferred according to average linkage algorithm based on simple genetic distances amongst the major population groups [Mountain et al., 1995]

This data can also be used to construct a tree summarising the relationship amongst the major human groups (Fig. 5). As expected, this tree separates out Africans from non-Africans. Among the non-Africans, the Europeans, Chinese and Indians are almost equally close to each other, being a little more separated from other Asians and New Guineans. The Indian population of today might then be surmised to have been put together by many ebbs and flows of people over the huge Eurasian continent.

Gene analysis reveals people radiating out of the Middle East and the Orient

To assess the patterns of these ebbs and flows, Cavalli-Sforza et al. (1994) have examined the frequencies of 69 genes from 42 populations covering all of Asia. Any given population is then represented as a point in the 69-dimensional space. This information can be summarised with the help of a multivariate analysis technique called principal components (PCs). The first PC for Asia explains 35.1% of the total variation in the gene frequency; the second PC explains 17.7% of the variation (Figs. 6 and 7). Subsequent components explain relatively little. These two maps are most instructive. The first PC map suggests that genetic affinities amongst Asian populations decline with distance along an East–West axis. This is compatible with movements of people radially fanning out from the Middle East, although it could also result from a westward movement along a very broad front in eastern Asia. In a similar fashion the second PC is compatible with fanning out of people from Southeast Asia and China, although it could also result from a major movement originating in the northernmost reaches of Asia.

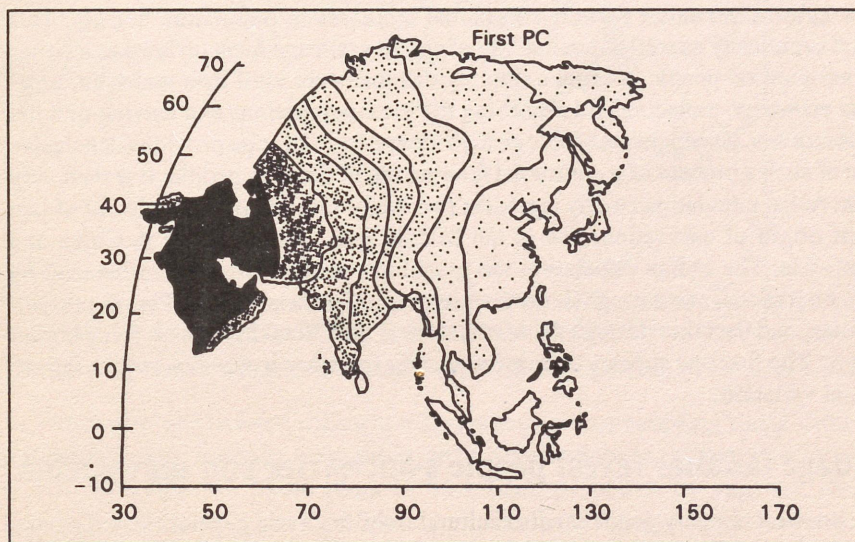


Fig. 6 Synthetic genetic map of the first principal component for Asia based on 69 genes for 42 populations [Cavalli-Sforza et al., 1994]

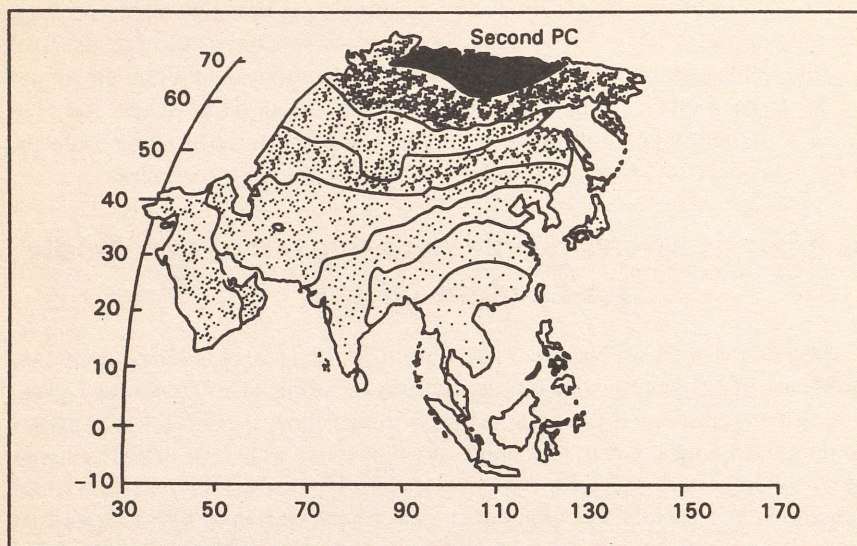


Fig. 7 Synthetic genetic map of the second principal component for Asia based on 69 genes for 42 populations [Cavalli-Sforza et al., 1994]

In both these cases, the first explanation, namely fanning out of people from the Middle East and from China and Southeast Asia is far more likely. These are known to have been two independent centres of origin of cultivated plants, the Middle East being the earliest in the world around 10 kybp, and China and Southeast Asia a little later around 8 kybp. Cultivation permits substantial increases in population density. This numerical superiority as well as availability of stored grain and meat on hoof as a buffer permits agricultural people to expand into regions that were until then under hunting-gathering economy, replacing and absorbing the local populations and leaving definite genetic footprints. Excellent archaeological evidence from Europe provides conclusive evidence of such a process of a northward fanning out of farming people. It is then very likely that Asian populations today represent two major radiations of people out of two centres of origin of cultivation, one in the Middle East, and the other in China and Southeast Asia. The Indian population too must have been profoundly influenced by these two migrations, one through its northwestern frontiers near Khyber Pass in present-day Pakistan, and the other through the northeast near the China-Myanmar-India border in Manipur. The first one appears to be more significant, since it explains twice as much of the total variation.

Language families reveal people's ancestries and movements

Humans transmit not only genes but also cultural traits from one generation to the next. Some of these are extremely conservative, being transmitted quite faithfully from parents to offspring. Foremost amongst these is language; children almost invariably acquire their mother tongue from their parents and other relatives. Language and other

conservative traits such as practices relating to disposal of the dead are therefore excellent devices to trace historical changes. If this be so, linguistic and genetic divergence ought to go hand in hand. To test this proposition, Cavalli-Sforza et al. (1994) plotted genetic distance amongst members of a human group against the number of different languages spoken by members of the group (Fig. 8). The excellent correlation confirms our faith in languages as good markers for unravelling the ancestries and movements of people.

The languages of the world have been classified into a number of major families. There are of course a few which are stand-alones, which cannot be assigned to any family. Nahali, a tribal language of Central India and Burushaski, spoken by a small group of people on the border of Pakistan and Afghanistan are two such. But all other languages of India can be assigned to one of four major language families—Austic, Dravidian, Indo-European and Sino-Tibetan. An excellent information base on the speakers of these languages is provided by the People of India Project of the Anthropological Survey of India. This project involved assigning the entire Indian population to 4635 ethnic communities, and putting together detailed information on each of them through interviews with over 25,000 individual informants spread over all districts of India, along with compiling information from a variety of published sources [Joshi et al., 1993]. This project records as the mother tongue the number of languages of different families spoken by Indian ethnic communities as indicated in Table 1.

Table 1

Language family	No. of languages in India	No. of Indian ethnic communities	Global distribution
Austic	24	65	Southeast Asia, Eastern and Central India
Dravidian	35	1281	South and Central India, Pakistan, Iran
Indo-European	110	3094	Europe, West Asia, North, West & East India
Sino-Tibetan	130	207	China, Southeast Asia, India bordering Himalayas

It is reasonable to assume that speakers of these four language families represent at least four major lineages [Parpola, 1974]. The first question to ask is whether these language families developed within the country, or came in with migrations of people from outside the subcontinent. The geographical range of distribution of Austic, Indo-European and Sino-Tibetan speakers is extensive; India harbours only a minority of the languages within these families. The geographical range of distribution of Dravidian languages is however restricted largely to India; there are only two outlying populations—Brahui in Baluchistan and Elamite in Iran. Dravidian languages might then have developed within India. Others are less likely to have done so, for we have no evidence of any major technological innovations that could have served to carry speakers of those languages outside India.

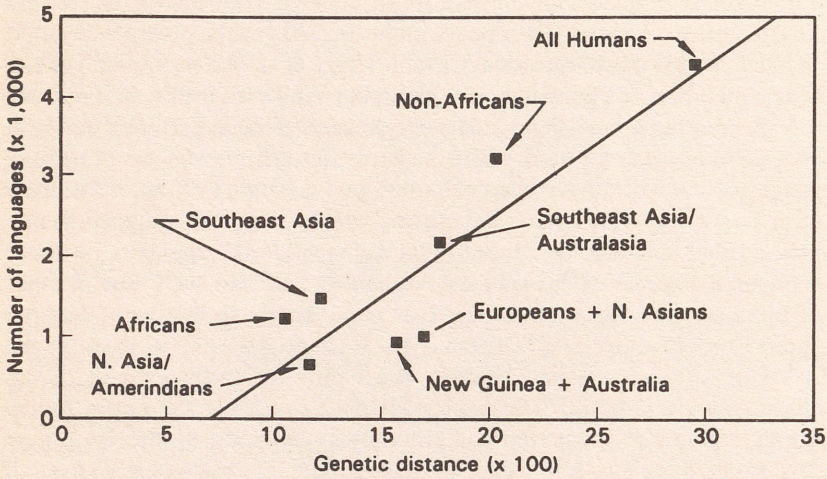


Fig. 8 Number of languages versus genetic distance in major regions of the world and for all of human populations [Cavalli-Sforza et al., 1994]

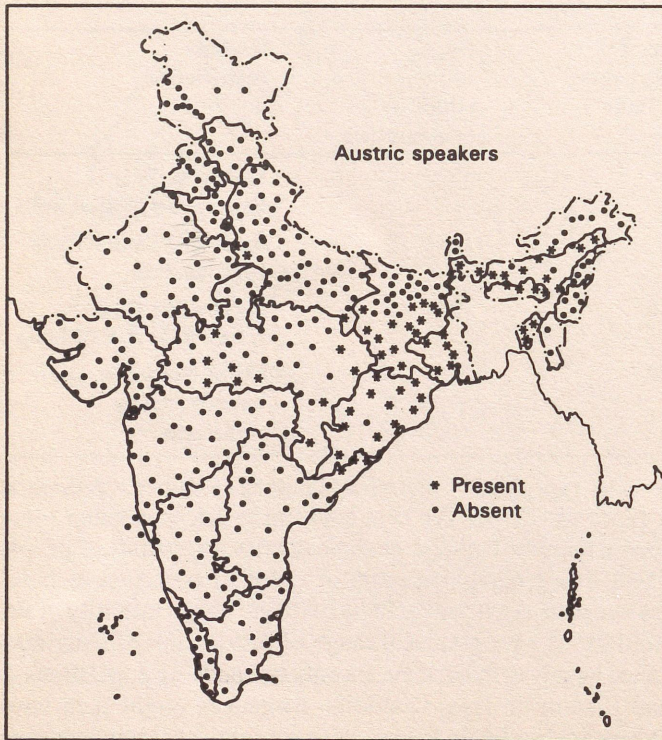


Fig. 9 District level distribution of Austric language speakers in India

Language and economy

We may look for evidence on how long the lineages belonging to different language families have been in India in two different ways. Firstly we may examine the current levels of economic activities of the communities speaking those languages and compare them with levels of economic activities of speakers of other language families. The tribal communities of India continue to extensively hunt and gather as well as practise low-input shifting cultivation. These communities are likely to have migrated to India relatively early, perhaps prior to the beginning of agriculture and animal husbandry. Some tribal groups or other speak languages belonging to each of the four families. Korkus, Mundas, Santals and Khasis speak Austric languages; Gonds and Oraons, Dravidian languages; Nagas and Kukis, Sino-Tibetan languages and Bhils and Varlis speak Indo-European languages (Figs. 9–12). But it is amongst Austric speakers that all communities are exclusively tribal. Outside India too, most Austric speaking communities practise very primitive technologies. This suggests that Austric speaking people may be the oldest inhabitants of India. They may be amongst the first group of *Homo sapiens* to have reached India, perhaps some 50–65 kybp. Since over 98% of Austric speakers today lie in Southeast Asia, they may have entered India from the northeast.

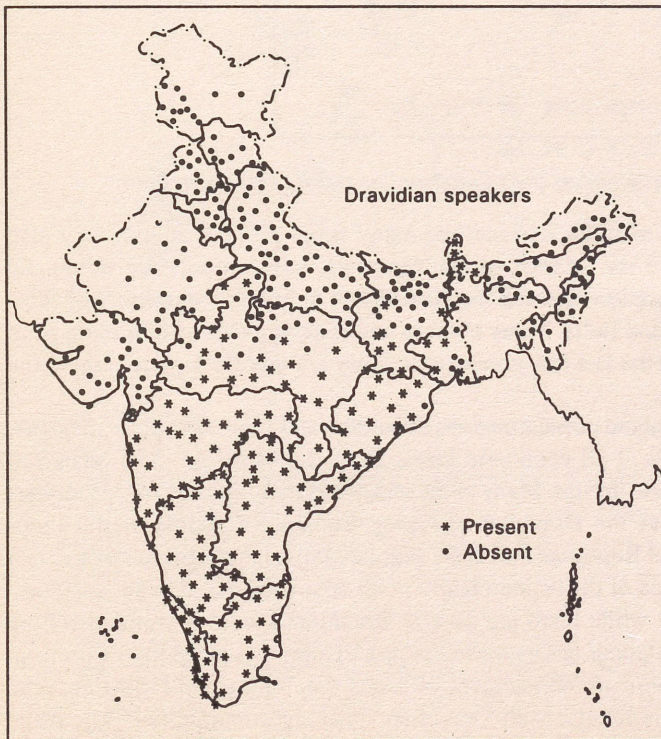


Fig. 10 District level distribution of Dravidian language speakers in India

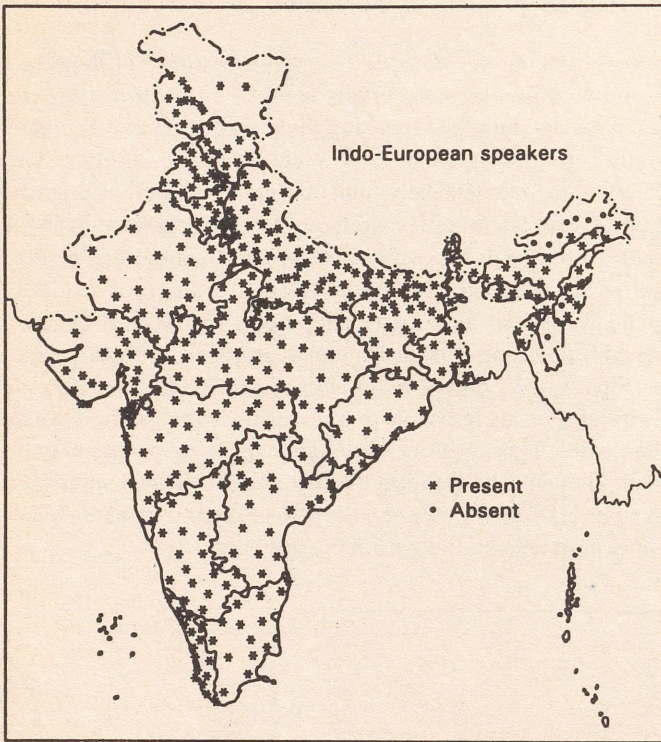


Fig. 11 District level distribution of Indo-European language speakers in India

Sino-Tibetan speakers of India also include many tribal groups, though they also include communities like the Maites of the Manipur valley practising advanced agriculture. Their concentration is along the Himalayas; only one community of West Bengal has reached mainland India. Many of them report having moved into India from Myanmar or China within the last few generations. They are therefore peripheral to the broader peopling of India.

The bulk of Indian mainland populations are Dravidian and Indo-European speakers. Both include communities at all economic levels from tribals to the most advanced cultivators, traders or pastoral groups. Many of the technologically less advanced amongst these communities such as the Dravidian speaking Kanians of Kerala or the Indo-European speaking Bhils of Rajasthan may have acquired these languages in more recent times through the influence of the economically more advanced mainstream societies. It is however notable that while there are several Dravidian speaking forest-dwelling tribal communities such as Gonds or Oraons in a matrix of more advanced Indo-European speaking communities, there are no enclaves of forest-dwelling tribal Indo-European speakers surrounded by more advanced Dravidian speaking communities. The tribal Indo-European speakers of South India are all nomadic communities such as Banjaras or Pardhis, with known history of migration from Rajasthan to South India in recent

centuries. This is strongly suggestive of Dravidians being older inhabitants of the Indian subcontinent, having been pushed southwards, surrounded by or converted to Indo-European languages by later arriving Indo-European speakers [Lal, 1974; Rakshit and Hirendra, 1980].

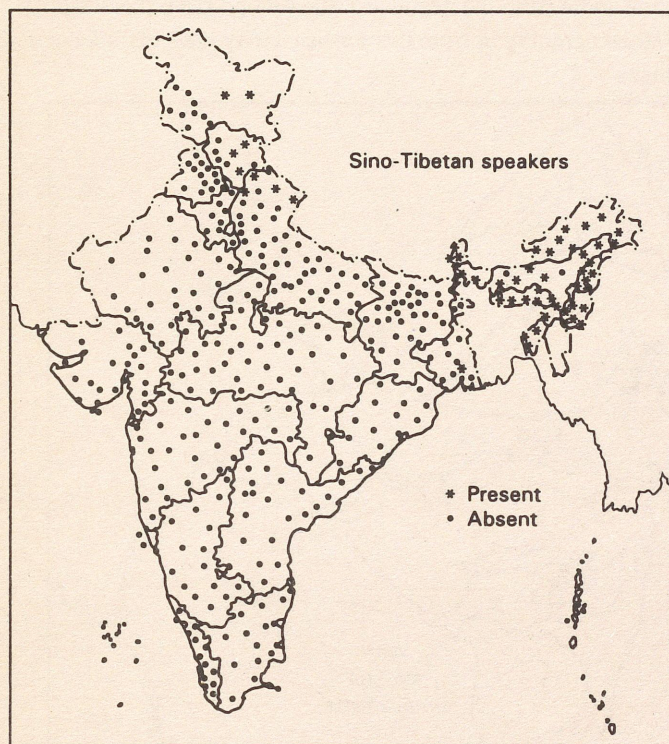


Fig. 12 District level distribution of Sino-Tibetan language speakers in India

One may then suggest the following sequence of migrations of these major language speaking groups into India: Austric–Dravidian–Indo-European. If this be correct, another interesting prediction follows. Austric languages having arrived in India earliest may show the most diversified vocabulary, Indo-European languages the least. To test this we have compiled words for universally used nouns such as mother, water, tree, etc. in several Austric, Dravidian, Indo-European and Sino-Tibetan languages. While a more objective analysis of the extent of such variation is under way, it appears that Austric languages show the greatest and Indo-European the least divergence.

Archaeological evidence

While tool-using *Homo erectus* populations have been in India for over 500 kyears BP, fossil human remains appear only after 45–50 kybp, associated with Middle Palaeolithic or Stone Age tools [Agarwal, 1982; Agarwal and Ghosh, 1973; Agarwal and

similar evidence of the two influences from northwest and northeast, with the western influence predominating over much of the country. Thus the Black and Red Ware reflect western, while the Cordedware reflects Chinese influence [Sankalia, 1974; Allchin and Allchin, 1988; Goudie, 1977; Brice, 1977; Rao, 1965, 1969].

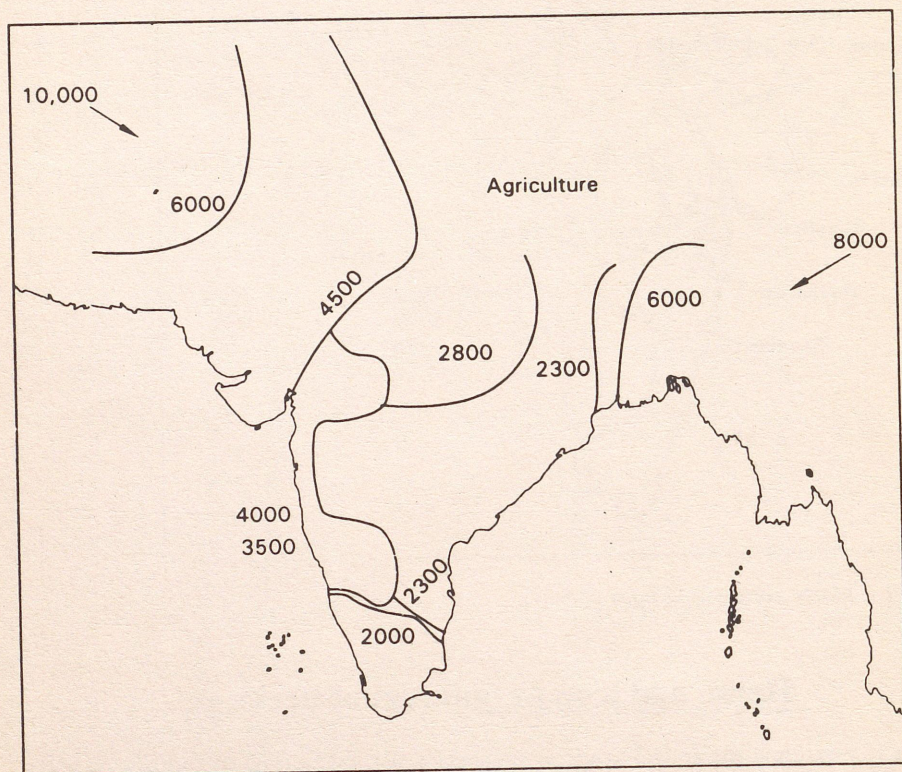


Fig. 14 *Contours of earliest dates of definite evidence of cultivation of crops in India*

It is likely that the farmers entering India from the northwestern passage were either Dravidian or Indo-European speakers; those entering the subcontinent from the northeastern passage may have been Sino-Tibetan or Austric speakers. If, as the linguistic evidence suggests, Dravidian speakers entered India well before Indo-European speakers then Mideastern farmers entering India from the northwest may have been Dravidian speakers. The remnants of related languages, Elamite and Brahui in Iran and Pakistan, are consistent with such a migration of Dravidian speakers from the Middle East to India.

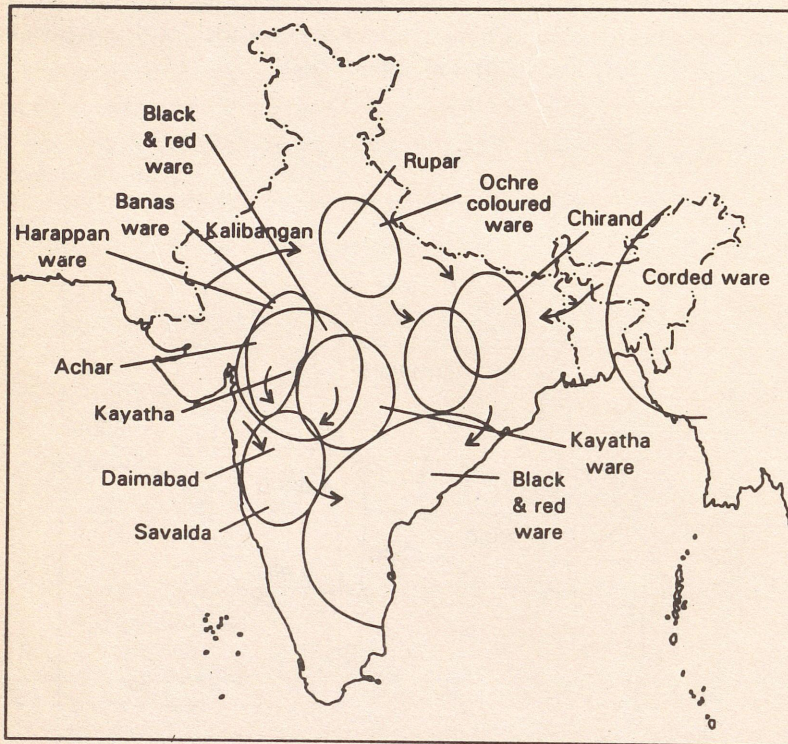


Fig. 15 Major pottery traditions of India

Horse and iron as pointers of heritage

If this is true, then the Indo-European speakers must have come to India with some other major advantage. Two technological innovations, known to have originated outside of India, are excellent candidates. They are the domestication of the horse, around 6 kybp on the shores of the Black Sea in present Ukraine, and the use of iron, around 5 kybp in Anatolia in present-day Turkey. Riding of horses or hitching them to carts greatly increases the mobility and the military or trading capabilities of a group. While cattle, sheep, goat and pigs were all domesticated in the Middle East around 10 kybp, the horse was domesticated 4000 years later in a separate centre in the Asian steppes. The most favoured theory of the spread of Indo-European languages today is that it was the language of these horse people who came to dominate Europe, West Asia and much of India over the next 4000 years. As a ruling class, they are believed to have imposed their language over Europe, without making any major genetic contributions to the populations. They may have wielded parallel influence in India.

The horse appears in archaeological records between 2000 to 500 years after the first appearance of cultivation of crops and husbanding of cattle, sheep, goat and pigs in different parts of India (Fig. 16). Particular styles of burial appear to accompany the horse people. These burial styles show links with styles noted in the Central Asian homeland of Indo-European speakers, strengthening our belief in the possibility that the Indo-European speakers indeed made their way to India propelled by the advantage that the control over horses conferred.

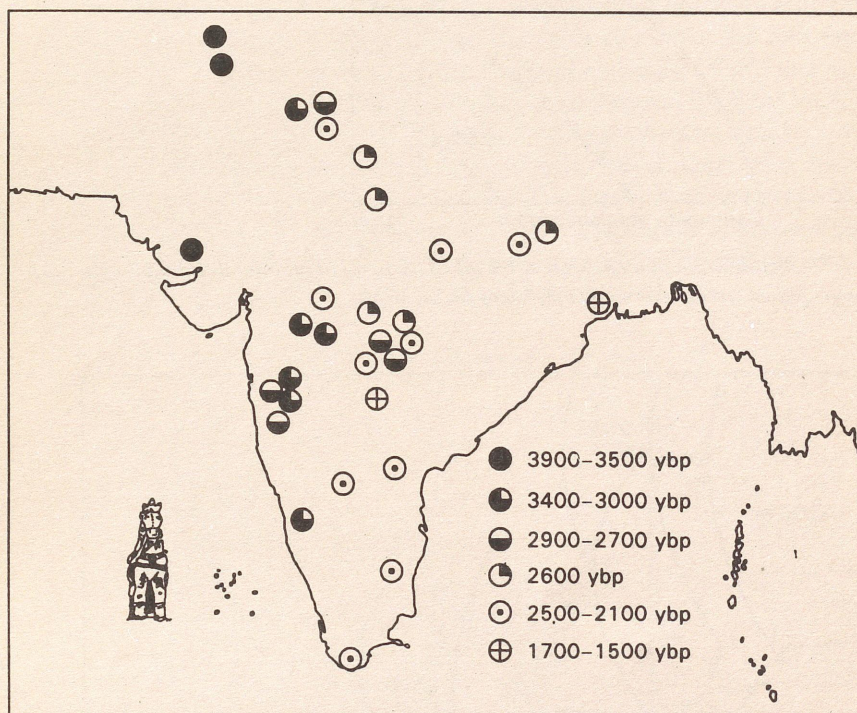


Fig. 16 *First dates of appearance of domesticated horse in different parts of India*

The people associated with Vedic traditions and Sanskrit definitely used horses, and may have been one group, though perhaps not the only group of Indo-European speakers to enter the subcontinent. These people also seem to have been associated with cremation as a method of disposal of the dead. Cremation is today the dominant mode amongst most Indo-European speaking communities of India while burial is common amongst Dravidian speaking communities, especially those affected little by the process of Sanskritisation (Fig. 17). This also suggests that Indo-European speakers came to India after Dravidian speakers, probably associated with the use of horses and the practice of cremation.

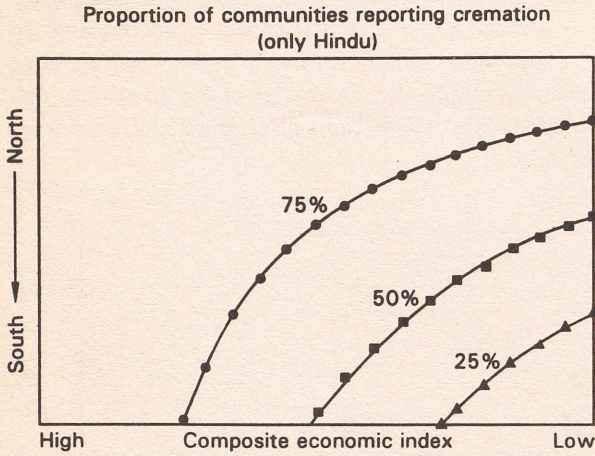


Fig. 17 The distribution of cremation vs burial of the dead in mainland India along a north-south geographical axis and amongst different social strata

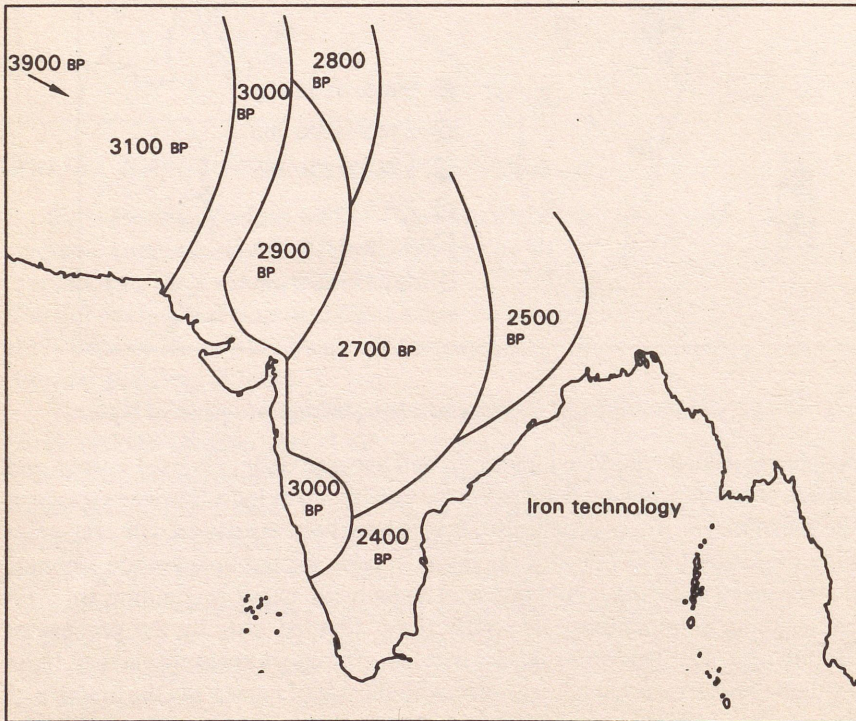


Fig. 18 The earliest known dates of appearance of iron technology in different parts of India

It is also possible that it was the use of iron that conferred an important advantage to certain groups of people migrating to India—groups that may have included speakers of Indo-European languages. The archaeological evidence suggests that use of iron is not necessarily associated with that of the horse, and appears either later than or ahead of the former in different parts of the country (Figs. 18 and 19). It is then likely that iron was brought to India by people other than the horse people, i.e., people other than Vedic people. Indeed, there may have been many waves of Indo-European speakers into India that may have brought into the country different languages of that family. Thus, some linguists believe that the present-day Indo-European languages came to India in at least two distinct streams, the first stream bringing in languages related to Bengali, Oriya, Marathi, Sindhi and Kashmiri, and the second stream of languages related to Punjabi, Hindi and Rajasthani [Gupta, 1979; Misra and Bagor, 1973; Goude, Allchin and Hegde; Agarwal, 1971; Chakrabathi, 1976; Banerjee, 1965; Stein, 1984; Jha, 1981; Parasher, 1992; Emmerick; Verma, 1971].

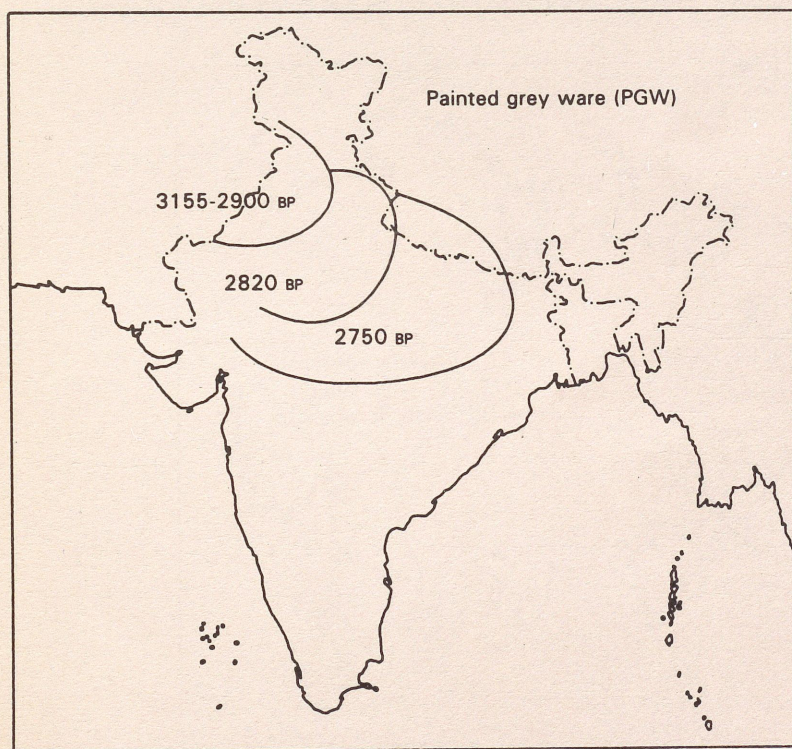


Fig. 19 *Known dates of appearance of painted grey ware pottery known to be associated with users of iron technology*

It is not at all clear whether the Harappan people spoke Dravidian or Indo-European languages. This civilisation is contemporaneous with the first appearance of the horse, most likely associated with Indo-European speakers in the archaeological record. It

could therefore have been an Indo-European speaking civilisation. But there is a greater possibility that it may have emerged out of the earlier Dravidian speaking communities of agriculturists. What seems more plausible is the equation of *Dasas* of Vedic people with the earliest, probably Austric speaking hunter-gatherers, and *Dasyus* with the Dravidian speaking cultivating communities. It is notable that the Vedic people were engaged in a violent conflict with *Dasyus*; such conflict may relate to struggle over fertile land [Possehl, 1979].

A plausible scenario

There are then many still unanswered questions pertaining to how our subcontinent was peopled. But the most plausible scenario is the one depicted in Figs. 20–23. The earliest of the modern *Homo sapiens* migrating into India, perhaps 50,000 years ago may have been speakers of Austric languages with the advantage conferred by the mastery over a symbolic language. Their genetic footprints may be discerned in the trends evident in

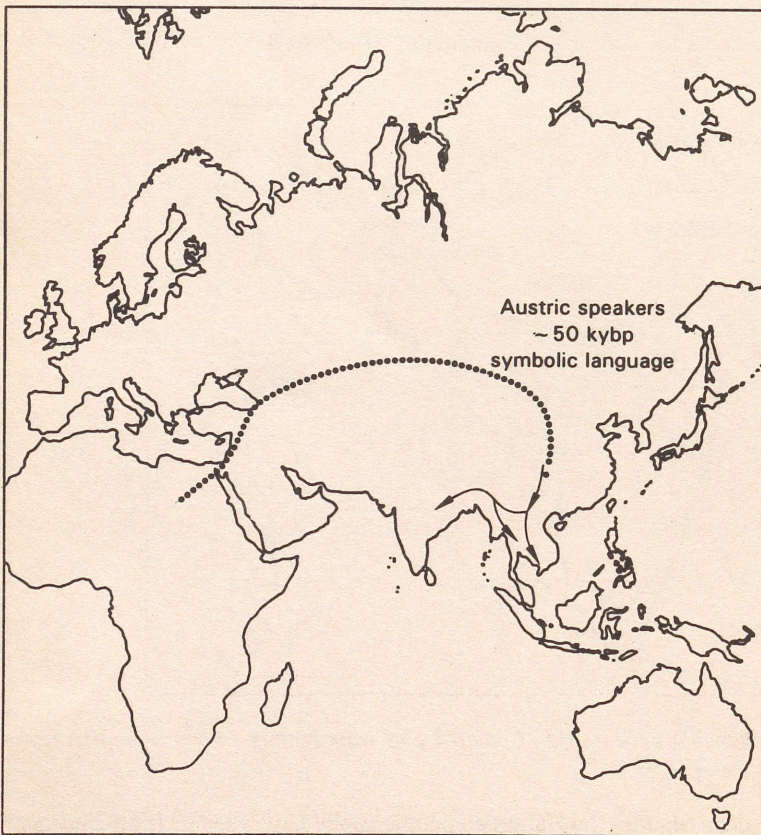


Fig. 20 A possible scenario of the earliest migrations of Austric speaking people into India

the second PC of the synthetic genetic map of Asia. The next major waves of migrations around 6 kybp may have been those of wheat cultivators from the Middle East and the rice cultivators from China and Southeast Asia. The former are likely to have been Dravidian speakers and probably contributed to the trend evident in the first PC of the synthetic genetic map of Asia [Allchin, 1963; Badam, 1984]; the latter may have been Sino-Tibetan speakers who would have contributed further to the trend revealed by the second PC. The latest major migration, around 4 kybp, may have included several waves of Indo-European speakers equipped with horses and iron technology.

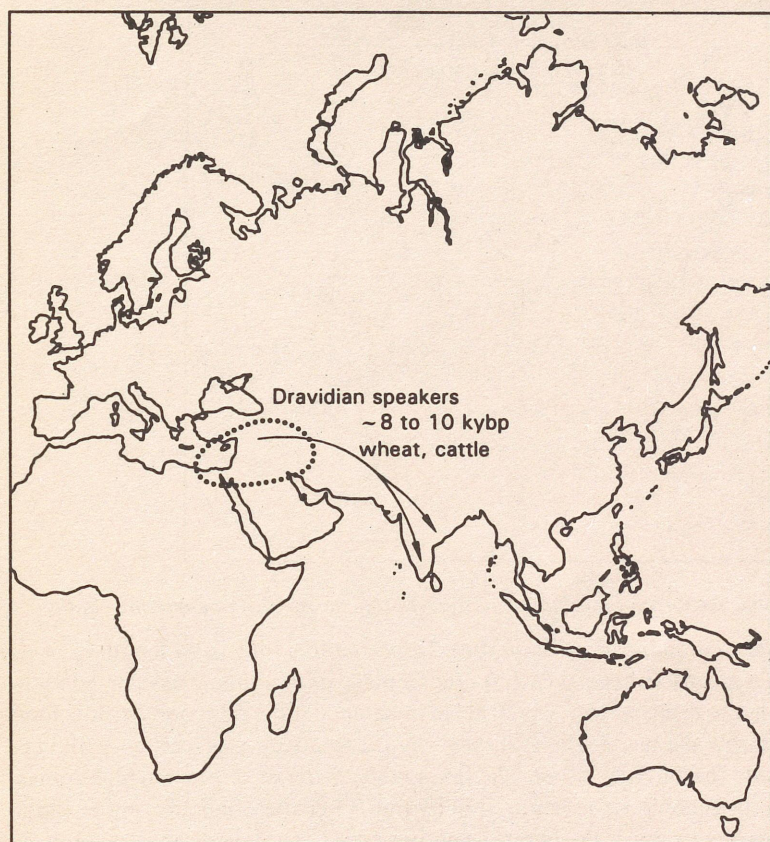


Fig. 21 *A possible scenario of migrations of Dravidian speaking people into India*

These might have been the most massive migrations peopling India. Others have followed, largely from the west, through the Khyber Pass on the northwestern frontier of the subcontinent. These seem to have been propelled by superior weaponry, increasingly better control over horses and, finally, sea-going ships. Such significant innovations may include some of the following. An important early development in weaponry was the composite angular bow which appeared in West Asia around 5 kybp. Bending through the length of the limb, releasing this bow string produced

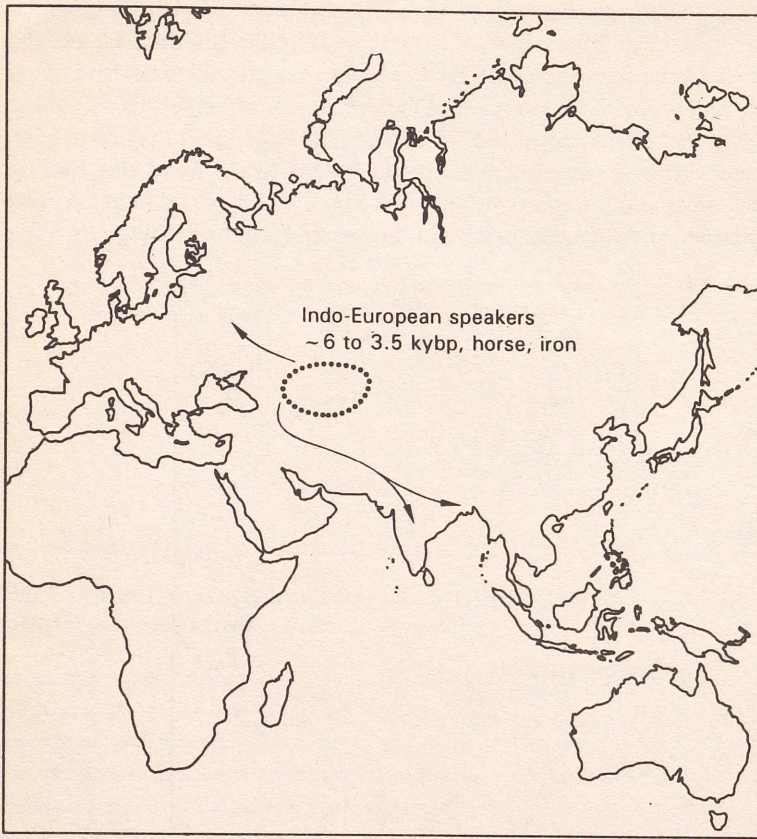


Fig. 22 A possible scenario of migrations of Indo-European speaking people into India

no kick, leading to a smooth and accurate shot. The extremely long draw length of over one metre led to a greatly enhanced cast. A crucial piece of equipment associated with control over the horse is the stirrup, which helps in balancing the rider and permits him to stand up and throw the lance. The earliest form of the stirrup was a string with two loops on either side for the rider's foot. The first known instance of iron stirrups comes from China in sixth century AD, reaching Iran by the 7th century and arriving in India with Turkish warriors in the 11th century. Another significant invention was the iron horseshoe, first known from Siberia in the 9th century AD, reaching India with Turkish warriors in the 13th century AD. Gunpowder was invented in China around AD 100 and slowly reached Iran, Arabia and finally Europe with the Mongols around AD 1400. It reached India with the arrival of the first Mughal emperor Babur, who used it in the First Battle of Panipat in AD 1526 (Fig. 24). The early cannons in India were made by welding together many iron rods. The Europeans introduced cast iron cannons in the next century which could fire more accurate and powerful volleys. The Europeans also developed superior ocean-going vessels from which cannons could be fired, by the 16th century [Deloche, 1983; Habib, 1992].

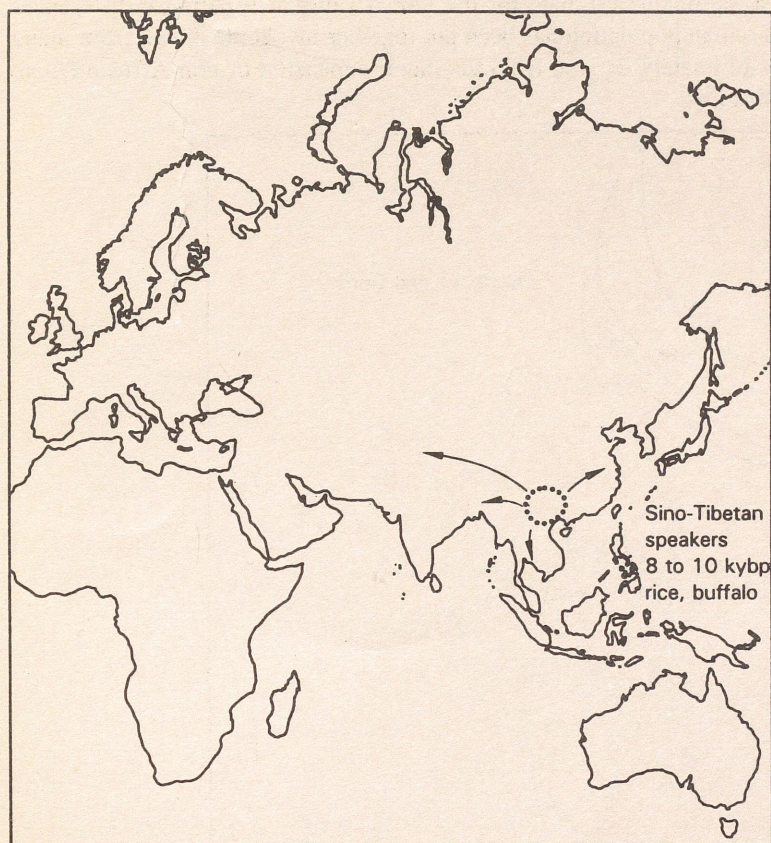


Fig. 23 A possible scenario of migrations of Sino-Tibetan speaking people into India

These many developments taking place in China, Central Asia and finally Europe brought in many people enjoying a military advantage (Fig. 25). The number of people thus coming in were probably not very large, but they contributed immensely to the cultural diversity of the country by bringing in new languages, new religions, and of course new technologies. Among these technologies was the spinning wheel, apparently invented in China and brought to Europe by the Mongols around the 12th century AD. It seems to have reached India in the 13th–14th century and created a tremendous commercial potential for textile production in India. Similarly, Indian agriculture too must have been greatly influenced by the introduction of the Persian wheel, first referred to by Babur in 1526–30 in his memoir. *Babur Nama*.

With these many streams of *Homo sapiens* coming into the country over 50,000 years or more, India has developed what Cavalli-Sforza calls an incredibly complex genetic landscape. Our mtDNA data on 101 Indians permits us to estimate the time to common ancestry of our people on the basis of the pairwise differences in the mtDNA sequences. These estimates of course depend on the assumed value of mutation rates; but

65,000 years is a reasonable estimate for the modal value of 9 pairwise differences (Fig. 26). So the Indian population has been put together by people drawn from many different streams ultimately derived from the major expansion of non-African *Homo sapiens* around this time.

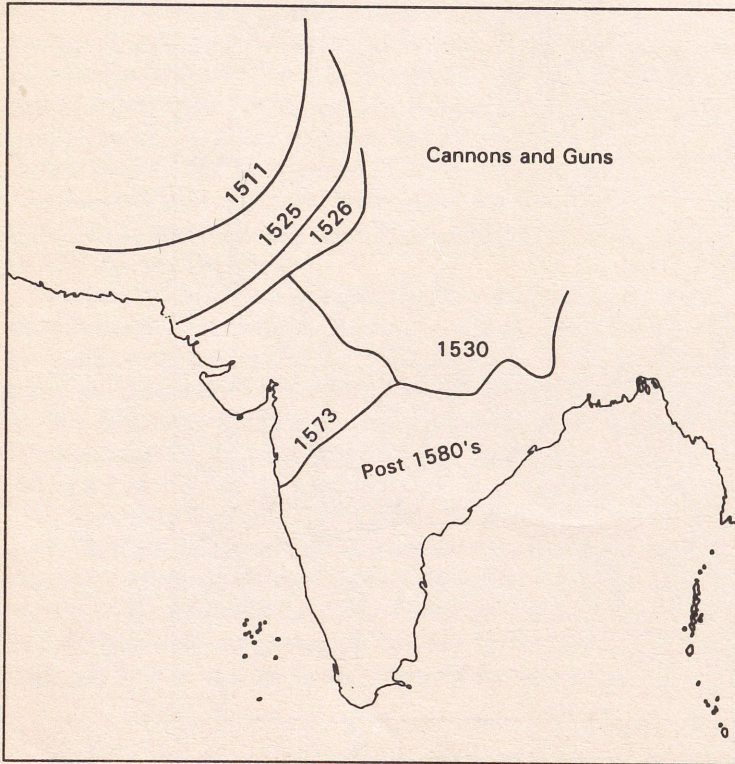


Fig. 24 The earliest known usage of cannons and guns in India

A segmented society

What the Indian population is remarkable for is the segmentation of this large population into thousands of endogamous groups. The data of the People of India Project recognises 4635 such ethnic communities. Many of these are however clusters of endogamous groups with similar traditional occupations and social status. The actual number of endogamous groups is decidedly much larger, of the order of 50,000 to 60,000 [Joshi, Gadgil and Patil, 1993; Gadgil and Malhotra, 1983]. This persistence of tribe-like endogamous groups, characteristic of hunting-gathering-shifting cultivation stage all over the world, in a complex agrarian and now industrial society of India, is a unique phenomenon. It seems to be the result of a peculiarly Indian tradition of subjugation and isolation, rather than the worldwide practice of elimination or assimilation of subordinated communities by the dominant groups.

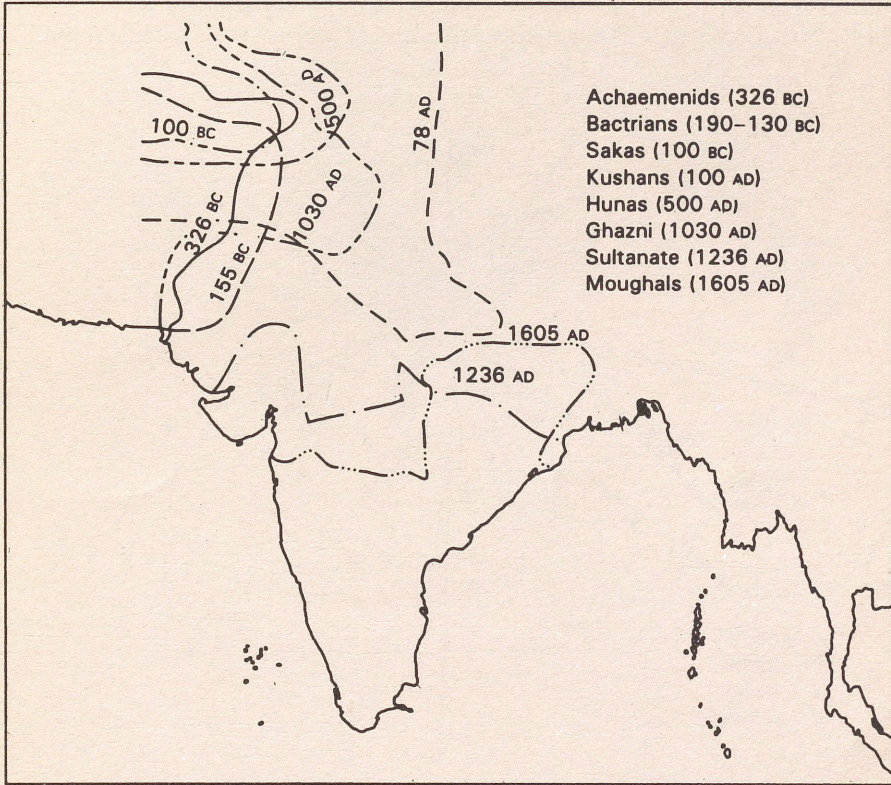


Fig. 25 *Migrations of people enjoying a military advantage through the northwestern passage into India between 300 BC and 1600 AD.*

Our mtDNA studies provide some notable insights into the structure of this social mosaic. For this purpose we chose two communities, Haviks and Mukris from the same district of Uttara Kannada. Haviks are a Brahmin group well known for their skills at growing multi-storeyed spice gardens of cardamom, pepper and betelnut. They also perform priestly functions, and are today prominent in many white-collar occupations. Their current population is around 100,000 individuals concentrated in an area of about 20,000 sq km. The Mukris, on the contrary, are members of a scheduled caste, earlier treated as untouchables. Their current population numbers around 9000 individuals concentrated in an area of 2000 sq km. They continue to indulge in substantial amounts of hunting, gathering and fishing to date, and serve as unskilled labour on Havik and other farms.

Figure 26 depicts the neighbour joining phylogenetic tree for 48 Haviks, 43 Mukris, 7 Kadars and a few other Indians. Note that Haviks and Mukris, although they lie at opposite ends of the social hierarchy, do not constitute two distinct trunks! Their sequences are intermingled, suggesting past genetic exchanges, although these may

have occurred well before the formation of the Indian caste system some 2000 years ago; indeed they may even derive from the time of common ancestry some 65,000 years ago, perhaps as a part of population expansion of non-Africans outside India.

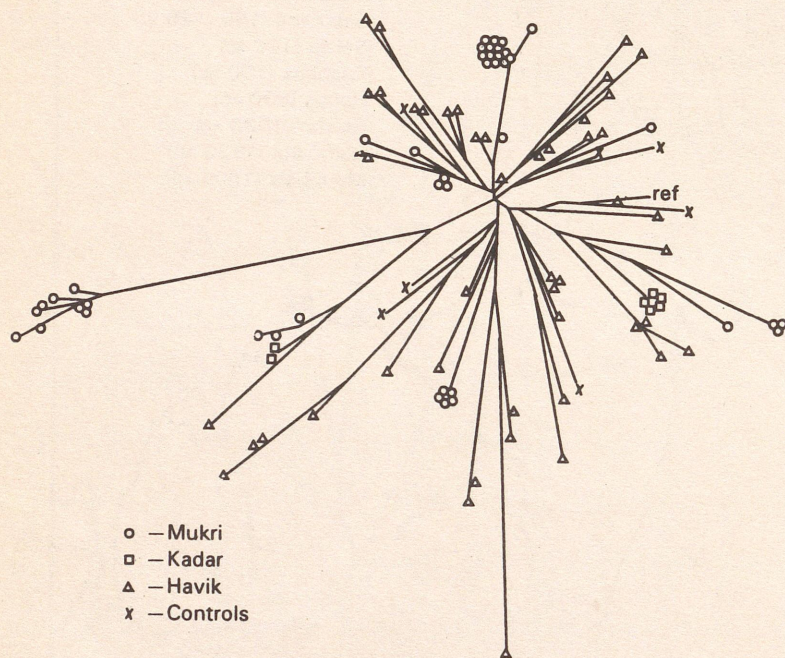


Fig. 26 Neighbour joining tree depicting 101 Indian sequences [Mountain et al., 1995]

But intermingled as they are, the Havik sequences form a distinctive star-like pattern with many short branches joining the centre, unlike the Mukri sequences which are bunched in a few clusters on long branches. The star-like Havik pattern is suggestive of a history of population expansion, while the clustered Mukri pattern suggests a long history of a stationary population, or a population that has experienced several bottle-necks. This is further brought out in the distribution of pairwise mtDNA base pair sequence differences for the Havik and Mukri populations (Fig. 27). The unimodal pattern for Haviks is compatible with a history of population growth while the multimodal Mukri pattern indicates a history of population stationarity or bottle-necks.

Such differences in genetic structure suggestive of different population histories have been suggested from other human populations earlier, but never before for two population groups living together in such a restricted geographical locality as a single district of Uttara Kannada. This reflects the unique history of Indian population, with dominant groups like Haviks enjoying high levels of resource access and expanding in numbers and range, while subjugated populations like Mukris existed side by side with much more limited resource access and stagnant populations (Fig. 28).

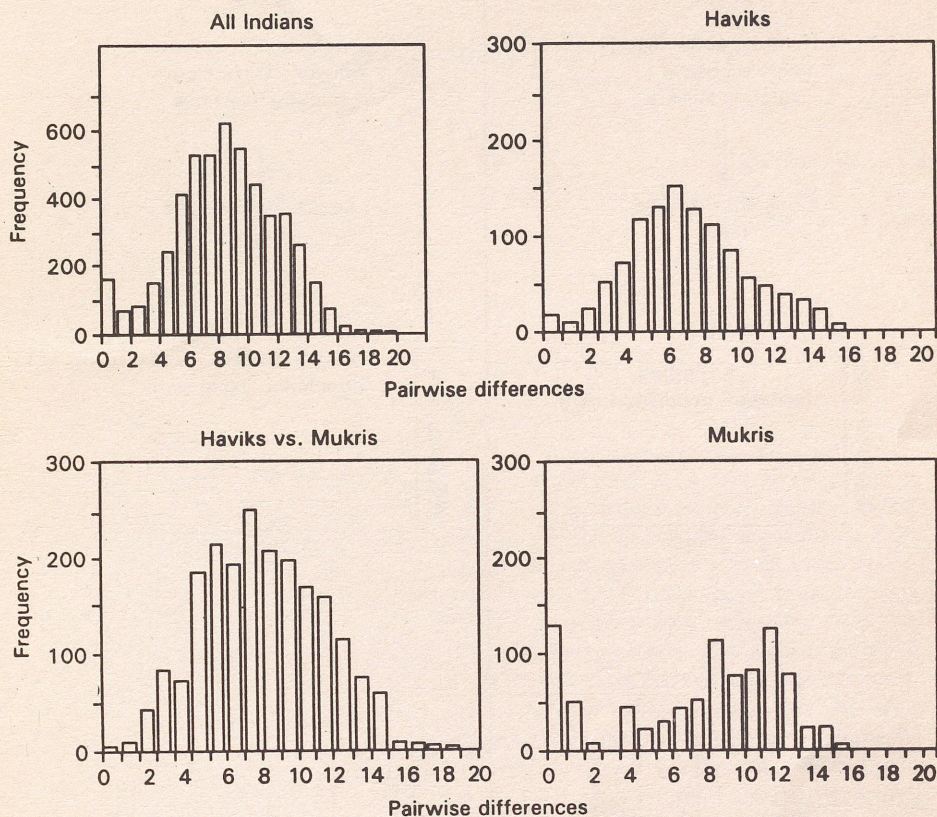


Fig. 27 *Distribution of pairwise differences in mitochondrial DNA sequences of 101 Indians [Mountain et al., 1995]*

Such scenarios have probably characterised the Indian social mosaic for long, perhaps since the beginning of cultivation and animal husbandry 6000 years ago. As groups with technologies conferring superiority in resource appropriation have migrated into and spread throughout India, they have subjugated other groups, restricted their resource access and permitted their continued existence, while the dominant groups have themselves grown in number and expanded in geographical range, perhaps dividing further into more endogamous groups.

This process of maintenance of a large number of communities in isolation from each other has been accompanied by extreme specialisation of occupation. It is perhaps this specialisation of occupation that has prevented Indians from cross-fertilisation of ideas and innovations, so that the Indian society has always been at the receiving end of technological innovations.

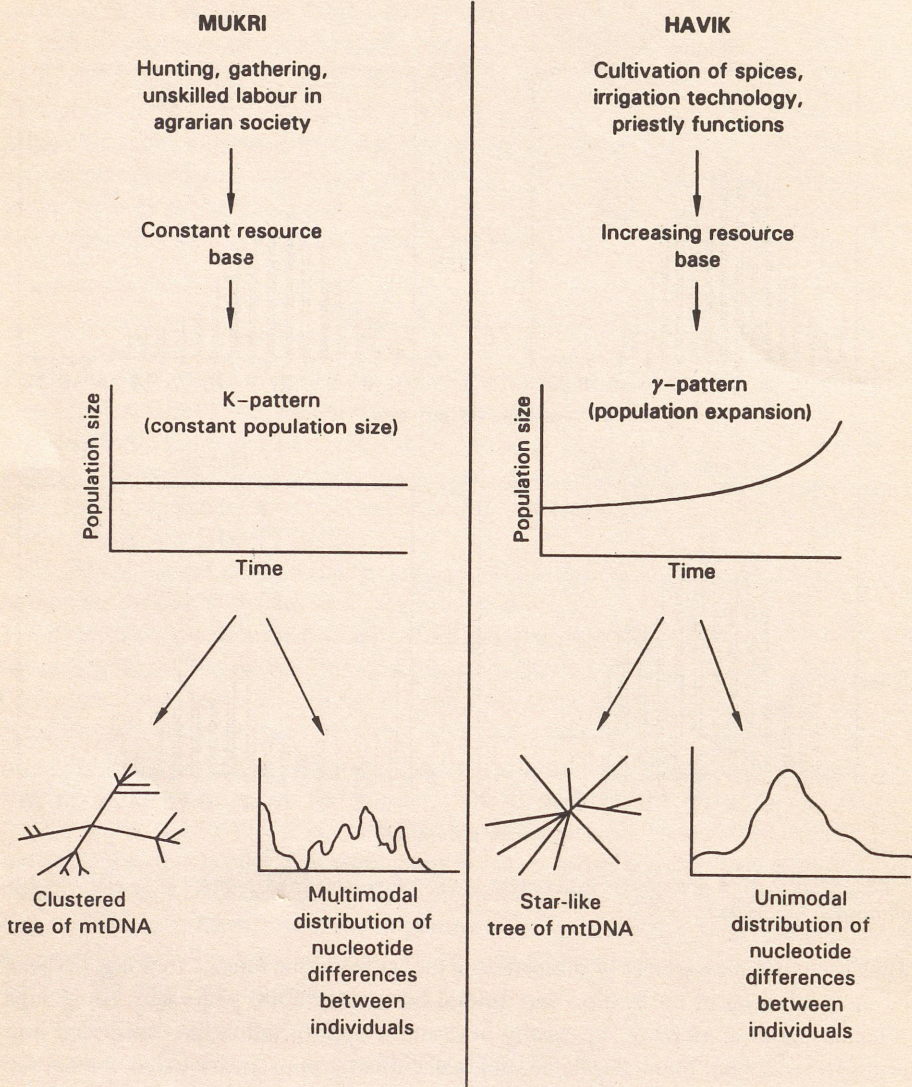


Fig. 28 Possible scenarios leading to the difference between the trees and pairwise difference distributions of the Mukri and Havik samples

Acknowledgements

We are grateful to our co-investigators in the population genetic study, Joanna L. Mountain, Joan M. Herbert, Peter A. Underhill, Chris Ottolenghi, L.L. Cavalli-Sforza and Silanjan Bhattacharya. We were greatly helped in the field work involved by Anindya

Sinha, A. Maithili and Vijayakumari of the Dr. Baliga College of Arts and Science, Kumta. Prema Iyer provided valuable assistance in investigating subsistence strategies of Haviks and Mukris in the field. M.D. Subash Chandran, K.M. Hegde, Nagu Mukri, Masti Mukri and many other friends from Uttara Kannada have also greatly helped us understand this society and its history. We are grateful to Dr. K.S. Singh and his colleagues in the Anthropological Survey of India for making the People of India dataset available, and for many stimulating discussions. This work has been supported by grants from the Ministry of Environment and Forests, Government of India, and from the Anthropological Survey of India.

References

1. Agarwal, D.P. (1971). *The Copper Bronze Age in India*. Delhi.
2. Agarwal, D.P. (1982). *The Archaeology of India*. London and Malmo.
3. Agarwal, D.P. and Ghosh, A. (1973). *Radiocarbon and Indian Archaeology*. Bombay.
4. Agarwal, D.P. and Kusumgar, S. (1974). *Prehistoric Chronology and Radiocarbon dating in India*. New Delhi.
5. Agarwal, D.P. and Pande, B.M. (ed.) (1977). *Ecology and Archaeology of Western India*. Delhi.
6. Allchin, B. and Allchin, R. (1988). *The Rise of Civilization in India and Pakistan*. Cambridge World Archaeology. Cambridge University Press, Cambridge.
7. Allchin, F.R. (1963). *Neolithic Cattle-keepers of South India*. Cambridge.
8. Badam G.L. (1984). Holocene faunal material from India with special reference to domesticated animals. In Juliet Clutton-Brock and Caroline Grigson (eds.). *Animals and Archaeology: Early Herders and their Flocks*. Oxford. 339–353.
9. Ballinger S.W., Schurr, T.G., Torroni, A., Gan, Y.Y., Hodge, J.A., Hassan, K., Chen, K.H., et.al. (1992). Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. *Genetics*. **130**, 139–152.
10. Banerjee, N.R. (1965). *The Iron Age in India*. Delhi.
11. Brice, W.C. (ed.) (1977). *Environmental History of the Near and Middle East*. London.
12. Cavalli-Sforza, L.L., Menozzi, P. and Piazza, A. (1994). *The History and Geography of Human Genes*, pp. 541. Princeton University Press, Princeton, and *Genetic Maps*, 518.
13. Chakrabarti, D.K. (1976). The beginning of Iron in India. *Antiquity*. **50**, 114–24.
14. Dani, A.H. et al. (1967). Timargarh and the Gandhara grave culture. *Ancient Pakistan*. **3**, 407.
15. Deloche, J. (1983). Geographical considerations in the localisation of ancient sea-ports of India. *The Indian Economic and Social History Review*. **20**, (4) 439–448.
16. Emmerick, R.E. Indo-Iranian Languages. *Encyclopaedia Britannica*. **9**, 438–57.
17. Gadgil, M. and Guha, R. (1992). *This Fissured Land: An Ecological History of India*. Oxford University Press, New Delhi and University of California Press, Berkeley.
18. Gadgil, M. and Malhotra, K.C. (1983). Adaptive significance of the Indian caste system: an ecological perspective. *Annals of Human Biology*. **10**, 465–478.
19. Goude, A., Allchin, B. and Hegde, K.T.M. Former extensions of the Great Indian Sand Desert. *The Geographical Journal*. **139**, 243–57.
20. Goudie, A.S. (1977). *Environmental Change*. Oxford.
21. Gupta, S.P. (1979). *Archaeology of Soviet Central Asia and the Indian Borderlands*. Delhi. **1**.
22. Habib, I. (1992). Pursuing the history of Indian technology: Pre-modern modes of transmission of power. *Social Scientist*. **20**, (2&3), 22.

23. Jarrige, J.F. and Lechevallier, M. (1973). Excavations at Mehrgarh. In M. Taddei (ed.) *South Asian Archaeology*. 77, 463-535.
24. Jha, D.N. (1981). Relevance of Peasant state and Society to Pallava-Chola times. *The Indian Historical Review*. 8, (1&2), 74-94.
25. Joshi, N.V., Gadgil, M. and Patil, S. (1993). Exploring cultural diversity of the people of India. *Current Science*. 64, (1), 10-17.
26. Kennedy, K.A.R. (1980). Prehistoric skeletal records of man in South Asia. *Annual Review of Anthropology*. 9, 391-432.
27. Lal, P. (1974). The tribal man in India: A study in the ecology of the primitive communities. In M.S. Mani (ed.) *Ecology and Biogeography in India*. The Hague. 281-329.
28. Megaw, J.V.S. (1977). *Hunters, Gatherers and First Farmers Beyond Europe*. Leicester.
29. Misra, V.N. and Bagor (1973). Late Mesolithic settlement in northwest India. *World Archaeology*, 5, 92-110.
30. Mountain, J.L., Hebert, J.M., Bhattacharyya, S., Underhill, P.A., Ottolenghi, C., Gadgil, M. and Cavalli-Sforza, L.L. (1995). Demographic history of India and mtDNA-sequence diversity *Am. J. Hum. Genet.* 56, 979-992.
31. Parasher, A. (1992). Nature of society and civilization in early Deccan. *Indian Economic and Social History Review*. 29, (4), 437-477.
32. Parpola, A. (1974). On the protohistory of the Indian languages in the light of archaeological, linguistic and religious evidence: an attempt at integration. In Van Lohuizen (ed.) *South Asian Archaeology*, 73.
33. Possehl, G.L. (1979). *Ancient Cities of Indus*. New Delhi.
34. Rakshit, Hirendra K. (1980). Ethnohistory of the tribal population of middle India. *Journal of Indian Anthropological Society*. 15, 97-112.
35. Rao, N.M.S. (1965). *The Stone Age Hill Dwellers of Tekkalakota*. Ponna.
36. Rao, N.M.S. (1969). *Excavations at Sangankollu*. Ponna.
37. Sankalia, H.D. (1974). *Prehistory and Protohistory of India and Pakistan*. 2nd edn. Ponna.
38. Stein, Burton. (1984). Politics, Peasants and the deconstruction of feudalism in medieval India. *The Journal of Peasant Studies*. 12, (2&3), 54-86.
39. Verma, B.S. (1971). Excavations at Chirand. *Puratattava*. 4, 19-23.
40. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. and Wilson, A.C. (1991). African populations and the evolution of human mitochondrial DNA. *Science*. 253, 1503-1507.
41. Vishnu-Mittre. (1977). India: local and introduced crops. In J. Hutchinson, G. Clark, E.M. Jope and R. Riley (eds.). *The Early History of Agriculture*. London.
42. Vishnu-Mittre. (1989). Forty years of Archaeobotanical research in South Asia. *Man and Environment*. XIV, (1), 1-16.