

# ***HISTORY AND GEOGRAPHY OF HUMAN GENES***

***A. PIAZZA***

***Dipartimento di Genetica, Biologia e Chimica Medica  
Università di Torino, via Santena 19, 10126 Torino, Italy***

## ***1. Introduction***

The collaboration between the Forensic Community and the scholars of human population genetics, among whom I place myself, has always been fruitful and of reciprocal benefit in Italy as much as in Europe and in North-America, in the latter with a more dialectical attitude as shown by recent rather hot debates. DNA analysis is today offering new possibilities of collaboration. Case work and search for reference populations complement in the daily activity of the forensic scholars. Substantial DNA databases have already been established for a number of population groups, but the development of new standards and new reference databases is likely in the immediate future following the implementation of PCR-based DNA typing systems, and a strong argument can be made for population geneticists to share protocols and markers with the forensic community in order to type well-defined reference populations and from them to contribute to the analysis of human genetic diversity. My talk, however, does not address to the future, but rather to the past: I have had the chance to analyze, over the past ten years, many genetic data from human populations and I am going to give a short review of our recent analyses. Our main interest lies in their interpretation in terms of prehistory and history of our species: a more comprehensive treatment will appear in a forthcoming book written in collaboration with Cavalli-Sforza and Menozzi [1].

## ***2. Expansions in Paleolithic***

A large fraction of the earth was settled by modern humans already a long time ago. At the end of Paleolithic, about 10,000 years before present, the number of inhabitants of the Old World was somewhere between one and ten million inhabitants [2]. We believe major genetic differences had already evolved in Paleolithic and that the expansions which took place at that time left a genetic trace which has not been completely erased by later events. Population

densities were especially low, and the effect of genetic drift was therefore quite extreme, while later numerical increases progressively reduced the importance of drift and genetic differentiation. We should like to elaborate on this idea and to show how these expansions "punctuated" the evolution of human populations [3].

There were probably at least two major expansions of the genus *Homo*. Starting between one and two million years ago, *Homo erectus* spread from Africa to all of the Old World. Anatomically modern *Homo sapiens sapiens* is believed by a number of paleoanthropologists to have started spreading from Africa between 130,000 and 60,000 years ago and to have continued moving until the present. The finding of old and reasonably well-dated modern humans in West Asia may somewhat change the origin of this second *Out of Africa* expansion, but archaic *Homo sapiens* closest to modern humans seems to come from Africa [4]. Other paleoanthropologists [5], however, prefer a hypothesis of local development of differences based on local continuity of East Asian skull morphology.

The stimuli to these major radiations from Africa or less probably from East Asia must have been either cultural or biological developments, or both, connected directly or indirectly with the efficiency of food collection and/or with drives or incentives to expand. The first radiation from Africa is that of humans with a larger brain and a relatively sophisticated set of stone tools. It seems obvious that an improved tool arsenal gave considerable advantage in the yield of food collection, and the chances and opportunities of spread to new habitats. A major coevolution of culture and perhaps brain biology has probably anticipated the origin of modern humans and their radiation. The human brain had reached its present size already before the second expansion, with the appearance of *Homo sapiens*. It is also likely that during this period modern human language went through the last stages of its evolution to its present degree of sophistication, and was the major support of the second expansion. While this is speculative, there is hard evidence that in the period between 100,000 and 60,000 years ago, prior to the expansion or accompanying its beginning, there was a substantial modification of the tool arsenal [6]. In addition, local cultures showed increased differentiation of stone tools, as if an incipient linguistic differentiation favored partial segregation and divergence of rapidly evolving local technologies [7]. Use of rafts, boats etc. must have been common because around 55 or 60,000 years ago several sea tracts had to be passed for reaching Australia [8].

One cannot exclude that navigation may have helped Africans to reach S.E.Asia (and finally Australia) along the southern coasts of Asia. Great advantage must have been provided by a fast and precise system of verbal communication while expanding to new territories and environmental niches, involving quite different climatic conditions not to mention flora and fauna.

### ***3. Synthetic genetic maps of the world***

Population movements after Paleolithic did not have the time to completely erase the differences generated in Paleolithic, but added new genetic patterns still recognizable today. No single gene would allow to clearly trace changes operated by later migrations on the original differences, but combining the information from as many genes as possible allows such reconstructions. We introduced the methodology of "synthetic" genetic maps, which are geographic maps of lines of equal values of principal component values (PCs) calculated from many gene frequencies [9]. This map generates a "genetic landscape" which is independent of all the others, and explains a fraction of the total genetic variation. We have shown by simulation that different PCs can separate expansions which started at different times and places [10]. The genetic variation explained by a PC is a function of the magnitude of genetic differences between the expanding population that migrates to a new area, and the aboriginal populations.

The synthetic map of the world [1] is in excellent agreement with the hypothesis of an expansion of modern humans from Africa: an extreme PC value is in Africa, the other in Australia. The synthetic map cannot choose between these two poles as origin of the expansion, but the archaeological records shows Africa was occupied by modern humans (and their ancestors) long before Australia. It also shows that much of human variation is in the west-east direction and therefore that adaptation to different climates cannot be the major determinant of variation observed at the genetic level.

### ***4. Neolithic and post-neolithic expansions***

The Neolithic transition is the passage from foraging (hunting and gathering), i.e. food collection, to food production by agriculture and animal breeding. Its earliest development was

in the Middle East, beginning around 10,000 years ago . Cereals and animals domesticated in the Middle East spread slowly in all directions shortly after 10,000 years ago. The expansion to Europe is best known and took place at the average rate of 1 km. per year [11]. The question was asked, if the spread from the area of origin was cultural (i.e., it was farming which diffused) or demic (farmers radiated) [12].

We originally developed synthetic maps [9] as a device to further test this hypothesis. If farmers truly spread from the Middle East in early Neolithic to all of Europe, it might be possible to detect still today traces of this migration in the genetic map of Europe. In particular, a comparison was made between the geographic map of the spread of agriculture to Europe, based on archeological dates of first arrival of farming in various regions of Europe [11] and the genetic maps. The correspondence between the archaeological map and the synthetic map corresponding to the first PC was remarkably clear [9]. On this evidence it was concluded that the diffusion of farming from the Middle East to Europe was a spread of farmers and not simply that of the technology. The correlation of the first PC values with the dates of first arrival of Neolithic farming (given in [11]) is 90%. Maps of Europe obtained with PCs of lower rank are also quite interesting. The second map of Europe shows a north to south gradient, which may reflect a climatic effect, but may also be at least in part of ethnic origin. In the north of Europe live Lapps, who have genetic similarity with other north-Eurasian populations from Siberia, and speak a language from the Uralic family, found today mostly east and west of the Urals. Lapps have undergone very substantial admixture with Scandinavians [13] and in spite of their probable west Siberian origin are now classified as Europeans, although they diverge considerably from all of them, being the extreme European outliers. Southern Scandinavians expanded into the area occupied in the north by people of Uralic origin like Lapps, establishing a gradient. Thus the north-south gradient may in part be due to selective effects of climate, and in part to contributions from East Asian populations related to the Uralic ones.

The third PC of Europe strongly suggests another possible expansion. It clearly defines a nuclear area between the Dniepr and Volga rivers, north-east of the Black Sea and north-west of the Caucasus. This is the same region where archaeological data place the beginnings of the "kurgan" culture, believed to have been developed, starting around 6,000 years ago, by pastoral nomads of the Eurasian steppes. Kurgan is the Russian name for a burial mound, very common

in the central Russian steppes. A connection was suggested between the kurgan culture and migrations of Indo-European speakers by Gimbutas [14].

Recently, Renfrew suggested that Indo-European languages came to Europe from Anatolia with the spread of farmers [15]. His rationale was that farmers did spread as discussed above, and they must have carried to Europe their original language. This must have undergone considerable evolution in the process of spread, but replaced almost all Pre-Indo-European languages spoken before in Europe, with one clear exception: Basque. Hypotheses on the place of origin of a language are notoriously difficult to support or disprove. Different authors suggested many different homelands for Indo-European [16] emphasizing the uncertainty associated with this decision. The problem is compounded by the possibility that there may have been several origins, scattered in time and space. The Gimbutas hypothesis of an origin in the kurgan region and spread during the Bronze Age (between 4,000 and 2,500 BC), which seems to have greatest support from archaeological and other considerations discussed in [17], is really not alternative to the Anatolian origin. Both hypotheses may be correct, the kurgan origin being much later, and secondary to the spread of agriculture from its origin to the steppes, and the establishment there of the new economy of pastoral nomadism.

The success of pastoral nomads was not limited to the expansion of the kurgan culture to Europe, but extended to a great number of other, some quite recent radiations (e.g. Turks in the last millennium). Unquestionably, however, the most spectacular of these secondary developments of agriculture was pastoral nomadism in the steppes of southeastern Europe and central Asia. The steppes extend from Romania to Manchuria in an almost uninterrupted way. The domestication of the horse took place in the western steppes, about 5,000 years ago or earlier, and was soon followed by military applications [18,19]. The way of life of pastoral nomads, their migrations in large numbers, demanded the development of considerable mobility, logistic and military skills. War chariots and later cavalry were important innovations which gave considerable military and political power to nomadic shepherds of central Asia. Nomads generated a hierarchical society in which they formed the highest castes, and the original farmers were organized in various lower castes distinguished by geography and profession. This social system was introduced by pastoral nomads from the Russian steppes in the Indian continent, where it survives today except where the Hindu religion never acquired control, or lost it to the Moslems.

Before leaving Europe for Asia, I should like to show you another relevant facet of the biology of Europe: what about anthropological traits?. More than a year ago, when the title of my talk was announced, I received a kind letter from a member of our Society expressing the hope that I would not restrict "genes" to "blood genes". In fact, the author of the letter - Dr. Røsing from the University of Ulm - is coauthor with Dr. Harding and Prof. Sokal of an interesting analysis of craniometric data collected by Dr. Røsing himself and Prof. Schwidetzky [20]. They consider 10 cranial measurements taken from fossils of three different periods: 1) before 3,000 B.C.; 2) 3,000-2,000 B.C.; 3) 2,000-1,000 B.C. The contour maps of their geographic variation in Europe show an East-West gradient very similar to that of our analysis of the third principal component of genes. The interpretation of this finding is not easy, even less so because of the low number of sites, around 50, on which the analysis is based. We can reasonably exclude, as the authors correctly do, an influence of the Neolithic spread of farmers on the cranial morphology, but two possible explanations remain open: a) a demic diffusion of genes (Indo-European?) from the Eurasian steppes as suggested by our genetic data or b) an unidentified influence by environmental factors which are known to affect craniometric and in general morphological traits more directly than genetic types.

The synthetic map of the first PC values of Asia shows a strong west to east gradient which has an apparent origin in the Middle East. This is likely to represent the eastern expansion of agriculture from the Middle East, which certainly reached as far as India in the eastern direction, and the steppes in the northern direction. The synthetic map of the second Asian PC values is an almost perfect gradient north to south. In the map of the third PC values there is an apparent center of expansion in the middle of Japan. In the last glacial period there was an almost complete continuity between Japan and continental Asia, both north and south, and the sea of Japan was practically an internal sea. Thanks to an archaeological survey of almost 28,000 sites [21], the curve of population growth in Japan has been reconstructed since the year 10,000. Around 4,000 B.P. the Japanese population reached its maximum of about 260,000 with a density of 0.89 per Km<sup>2</sup> and considerable variation from region to region. The period is called "Neolithic", although there was no agriculture until about 1,700 years ago. Foraging and fishing produced high population densities in Japan without agriculture. There is

a thus strong indication of an expansion west and north starting from Japan; the period is difficult to establish, but the date of population peak of 4,000 years ago may be a reasonable date for the expansion.

The genetic map of Africa is reasonably well-known, apart from the Sahara which is presently inhabited only in restricted zones. The synthetic map of the first PC values shows the distinction between the Caucasoid populations inhabiting the northern part of the Sahara, and black subsaharan populations. The Sahara was not a desert until a few thousand years ago, and was the seat of agricultural development, especially cattle breeding which may have been in part local. With the drying up of the Sahara, 3-4,000 years ago cattle had to migrate south and farmers (mostly dark-skinned) started domesticating plants like sorghum, millet and others in the Sahel region, immediately south of the Sahara. Much of these new agricultural developments were in West Africa, from where a well-known expansion began heading toward the east and south around 3,000 years ago or earlier. Its latest and most important part is usually called the Bantu expansion, because people at the boundary between Nigeria and Cameroon spoke Bantu languages and spread them over all of Central and most of South Africa. This radiation from West Africa is well visible on the synthetic map of the second PC values.

The synthetic maps of the Americas are more complicated, mostly because of extreme drift [22]. The synthetic map of the first PC values shows the approximate locations of three major groups of American natives: *Eskimos* in the extreme north, *Na-Dene* below them, and *American Indians* all over the rest. This classification was suggested by the linguistic work of Greenberg, odontological work of Turner, and preliminary genetic work of Zegura [23]. The synthetic map of the second PC values shows a heavy hybridization with European settlers in the east coast of North America, and the basic homogeneity of the Andean region. The analysis of the whole continent, however, fails to single out regions of spread of agriculture which, as well-known, started on the Mexican plateau and the northern Andes. Agriculture spread only lately from Mexico to North America, perhaps because it was difficult to cross the Sonoran desert in the north of Mexico, and the regions bordering Mexico further north were also very arid. It was only some 5,000 years after the beginnings of agriculture in Mexico that the first agricultural communities developed in the southern United States, and agriculture started spreading across the prairies [24].

### ***5. Major expansions and gene-culture evolution .***

The genetic map of the world seems to carry the traces of many human expansions. Each of them was most probably the result of some technological innovation which generated population growth and helped spreading people from the nuclear area of origin to new regions. Some of these may have been unoccupied or only more sparsely settled before, and especially in the latter case genetic clines must have formed.

Some major expansions occurred in Paleolithic and their traces are the cause of the major differences we still see among aboriginal populations of continents and parts of them. After Paleolithic, the major expansions were determined by the development of agriculture, which arose at very similar times in widely different regions. The earliest were in the Middle East , in China (several centers, one in the north and at least two in the south) and Mexico-northern Andes. Spread from the Middle East is very clearly visible towards Europe, towards Central and South Asia and to some extent towards North Africa. Arab and Bedouin migrations to North Africa in the seventh century AD may have made it more difficult to follow the initial spread. Similarly, genetic data do not yet do much to clarify the genetic effects of the Chinese agricultural expansions; but we know that north and south China have different ecology, neolithic history, genetic and cultural maps [25]. They grew at least initially different crops (millet in the north, rice in the south). Genetic data are not yet detailed enough to show much more than a sharp difference between north and south.

In summary, major expansions prior to the modern era provided a series of "punctuations" [3] due to cultural innovations which we consider major determinants of our genetic change. The spread of a population entails the spread of its culture: therefore it is simply natural that the first settlers carried their language with them: this was certainly true for the spread of Europeans after the great geographic discoveries. However, during long periods such as that of the spread of Neolithic farming to Europe which lasted nearly 4,000 years (from 9,500 to 5,500 years ago), the initially present languages would transform and differentiate sharply between regions and their differentiation is difficult to trace back, as, by definition, no written record of prehistoric events exist and prehistoric languages are inferred

very indirectly. In the past, less advanced societies, however, vertical cultural transmission (from parents to children) was common, and horizontal transmission (between unrelated people) rare, while it is very frequent in modern society [26]. Cultural traits transmitted vertically and therefore in a way more similar to that of genes must evolve like them, and one can expect to see coevolution of genetic and cultural traits thus transmitted, including languages.

## **6. Prospects for the future**

I would like to close this note by linking the soft *scenario* I just presented to the hard bench work of the forensic haematologist. A considerable part of his job is today that of comparing DNA profiles. More precisely major efforts are spent to calculate with as much precision as possible two likelihoods: 1) that of two matching DNA profiles coming from the same person; 2) that of two matching DNA profiles coming from two different individuals randomly drawn from a structured extant population. I would suggest that the former likelihood depends on the technology, i.e. the future: mini- and microsatellites, PCR-based typing, direct sequencing, etc., shall increase the degree of resolution by which matching of bands will be assessed or excluded with a more and more decreasing error. The latter likelihood is depending on how the population is genetically structured, i.e. on its past, its history, that history we try to interpret from their current genetic geographic images. The forensic scientist lies at the border: he looks at the future challenged by a moment of much innovative activity in DNA technology; but he must look back at the past as well in considering how populations evolved to the present structure. A way of not losing our head by this kind of squint is to share the challenge with scholars of other disciplines, and clearly I am thinking mainly of geneticists and statisticians. The study of human variation is one of the principal areas of interest for geneticists and the introduction of techniques that allow the direct study of DNA is going to add new dimensions to the analysis of human diversity. World-based and European projects have been proposed with the aims of planning systematic studies of genetic differences in suitably chosen samples of our species and saving, for future analyses, the DNA of a significant proportion of individuals representing current human variability. Many of the problems they are going to face are common to those of the forensic community, mainly the establishment of: a) standardized

laboratory protocols; b) centralized databanks; c) well-defined reference populations; d) common sets of DNA markers. It is time to join efforts. I know from my personal experience in Italy that the forensic scientists are very open-minded in this respect: I hope this Congress will be remembered as the one where the problem of identifying the suspect is paired also to the problem of identifying ourselves and our origins.

## REFERENCES

1. L.L.Cavalli-Sforza, P.Menzozi and A.Piazza, *The History and Geography of Human Genes* (Princeton University Press, Princeton, 1993). See also *Science* **259**, 639 (1993).
2. J.N.Biraben, *Population* **4**, 1 (1980); F.A.Hassan, *Demographic Archaeology* (Academic Press, New York, 1981)
3. N.Eldredge and S.J.Gould, in *Models in Paleobiology*, T.J.Schopf, Ed. (Freeman, San Francisco, 1972), pp. 82-115.
4. G.Brauer, in *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, P.Mellars and C.Stringer, Eds. (Princeton University Press, Princeton, 1989) pp. 123-154; P.G. Rightmire in *idem*, pp. 109-122; C.B.Stringer in *idem*, pp.232-244; R.G.Klein in *idem*, pp. 529-546.
5. F.Weindenreich; C.S. Coon, *The living Races of Man* (Knopf, New York, 1965); M.H.Wolpoff, in *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, P.Mellars and C.Stringer, Eds. (Princeton University Press, Princeton, 1989) pp. 62-108.
6. R.G.Klein, in *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, P.Mellars and C.Stringer, Eds. (Princeton University Press, Princeton, 1989) pp. 529-546.
7. G.L.Isaac, in *Origins and evolution of Language and Speech*, S.R. Arnald, H.D.Steklis and J.Lancaster, Eds. (New York Academy of Sciences, New York, 1976) vol. 280, pp. 275-288.
8. J.B. Birdsell, in *Sunda and Saul Prehistoric Studies in Southeast Asia, Melanesia and Australia*, J.Allen, J.Golson and R.Jones, Eds. (Academic Press, London) pp.113-167; R.M.Jones, in *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, P.Mellars and C.Stringer, Eds. (Princeton University Press, Princeton, 1989) pp.743-782.
9. P.Menzozi, A.Piazza, L.L. Cavalli-Sforza, *Science* **201**, 786 (1978).
10. S.Rendine, A.Piazza and L.L.Cavalli-Sforza, *American Nat.* **128**, 681 (1986).
11. A.J.Ammerman and L.L.Cavalli-Sforza, *Neolithic Transition and the Genetics of Populations in Europe* (Princeton University Press, Princeton, 1984).
12. A.J.Ammerman and L.L. Cavalli-Sforza, in *The Explanation of Culture Change*, C.Renfrew, Ed. (Duckworth, London, 1973), pp.343-357.
13. C.R. Guglielmino Matessi, A.Piazza, P.Menzozi, and L.L. Cavalli-Sforza, *Am. J. Phys. Anthropol.* **83**, 57 (1990).
14. M.Gimbutas, *The Civilization of the Goodness: The World of Old Europe*, (Harper, San Francisco, 1991),.
15. C.Renfrew, *Archaeology and Language: The Puzzle of Indo-European Origins*, (Jonathan London, 1987).
16. A.B.Dolgopolsky, in *Mediterranean Language Review*, A.Borg and P.Wexler, Eds., **3**, 7 (1988); M.Kaiser and V. Shevoroshkin, *Annual Review of Anthropology*, **17**, 309 (1988); T.V.Gramkeldidze and V.V.Ivanov *Sci.Am.* **262**, 110 (March 1990). p73
17. J.P.Mallory, In *Search of the Indo-Europeans: Language, Archaeology and Myth*, (Thames and Hudson, London, 1989).
18. M.Zvelebil, in *The Cambridge Encyclopedia of Archaeology*, A.Sherrat, Ed. (Crown, New York, 1980), chap.38;
19. J.M.Diamond *Nature* **350**, 275 (1991).
20. R.M.Harding, F.W.Rosing and R.R.Sokal. *Homo (Sonderheft Neolithikum)* **40,45** (1989, published 1990)
21. S.Koyama, *Prehistoric Japanese populations: a subsistence-demographic approach*, Unpublished.
22. D.H.O'Rourke, B.K.Suarez and J.D.Crouse, *Am.J.Phys.Anthropol.* **67**, 241 (1985); D.H.O'Rourke and B.K.Suarez, *Ann.Hum.Biol.* **13**, 13 (1986); F.M.Salzano and S.M.Callegari-Jacques, *South American Indians: a case study in evolution*, (Oxford University Press, Oxford, 1988).
23. J.H. Greenberg, G.C. Turner II and S.L. Zegura, *Current Anthropology* **27**, 477 (1986).
24. T.P.Culbert, in *Ancient South Americans*. J.D.Jennings, Ed. (Freeman, New York, 1983), chap.2.
25. N.Saitou, K.Tokunaga and K. Omoto, in *Isolation and Migration*, R.F.Roberts Ed. (Cambridge University Press, Cambridge, 1991).
26. L.L.Cavalli-Sforza and A.Feldman, *Cultural Trasmision and Evolution: a Quantitative Approach* (Princeton University Press, Princeton, 1981); B.S.Hewlett and L.L.Cavalli-Sforza, *Anthropology* **88**, 922 (1986).

# 1. Genetic typing and human rights

