# Demic Expansions and Human Evolution

Luigi L. Cavalli-Sforza, Paolo Menozzi, Alberto Piazza

**Geographic expansions are caused by successful innovations, biological or cultural, that favor local growth and movement. They have had a powerful effect in determining the present patterns of human genetic geography. Modern human populations expanded rapidly across the Earth in the last 100,000 years. At the end of the Paleolithic (10,000 years ago) only a few islands and other areas were unoccupied. The number of inhabitants was then about one thousand times smaller than it is now. Population densities were low throughout the Paleolithic, and random genetic drift was therefore especially effective. Major genetic differences between living human groups must have evolved at that time. Population growths that began afterward, especially with the spread of agriculture, progressively reduced the drift in population and the resulting genetic differentiation. Genetic traces of the expansions that these growths determined are still recognizable.**

Expansions are frequently the result of innovations, biological and technical, that increase the availability of food and thus determine local growth. Out-migration to relieve population pressure follows, and when this cycle repeats, expansion proceeds in all directions until unsurmountable barriers are reached. Innovations that favor transport and settlement in new environmental niches can also spark new expansions.

The radial rate of a population expansion depends on the growth rate of a population and the migration rate. Radial rates of 1 km per year have been measured, and there have probably been even faster expansions. Populations tend to differentiate genetically from neighbors because of divergence due to random genetic drift or natural selection to different environmental conditions. In turn, the expansions of populations spread their genes. Especially at times in the Paleolithic when population densities were low, drift was an important cause of local genetic differentiation, which led to a mosaic genetic geography. Expansions tend to add detectable patterns of nearly circular genetic gradients around their areas of origin and can extend to large regions in a few thousand years. Their genetic effects are relatively stable, and overlapping expansions that took place at different times can be distinguished from each other by statistical methods. Genetic geography is thus a powerful addition to prehistorical and historical analysis and generally reveals a number of recent expansions whose effects have not yet been completely erased. In a given area, one might anticipate phases of fast change due to arrival and settling of a population wave, followed by long, relatively quiet periods. This form of "punctuated" evolution is superficially similar to Eldredge and Gould's (*1*) speciation theory in showing major changes of gene frequencies over short periods of time. In this article, we summarize some recent, mostly unpublished observations of interest in the evolution of modern humans.

## Expansions in the Paleolithic

There probably were at least two major expansions of the genus *Homo*. Between 1 million and 2 million years ago, *Homo erectus* started to spread from Africa to the entire Old World. Anatomically modern *Homo sapiens* (*H. sapiens sapiens* or *H.s.s.*) is believed to have begun this spread between 100,000 and 60,000 years ago and to have continued moving until the present. The discovery of old and reasonably well-dated modern human remains in west Asia may change somewhat the proposed origin of this second expansion, but archaic *Homo sapiens* closest to modern humans seems to have come from Africa (*2, 3*). However, other researchers (*4*) have preferred a hypothesis of parallel local evolution in many continents that started with the first expansion from Africa of *H. erectus*. The major evidence for this notion is the claim of a continuity of East Asian skull morphology over a period of a million years or more.

The evidence from mitochondrial DNA (mtDNA) in favor of the expansion of *H.s.s.* has been somewhat weakened by a recent statistical reanalysis of the mtDNA tree roots (*5*). However, these criticisms will require further review and do not include other evidence from mtDNA, in particular the greater wealth of mtDNA types in Africa (*6*). Also neglected is the evidence from nuclear genes, which has recently been expanded and supports the notion that the spread of *H.s.s.* from Africa to Asia and from Asia to other continents agrees with the dates of first appearance of modern humans in the local paleoanthropological record (*7, 8*). However, current genetic data cannot entirely refute an intermediate hypothesis, namely, that humans who descended from their modern African ancestors expanded eastward and mixed to some extent with East Asians who were more direct descendants of archaic *H. sapiens*. This hypothesis might help explain both the observations on partial local continuity of East Asian forms and the recent expansion of *H.s.s.* from the west.

The stimuli to these major radiations from Africa must have been developments in culture, biology, or both. These changes were connected with the efficiency of food collection or of geographic expansion, as well as with incentives to expand. The first radiation from Africa is that of humans with a larger brain size than that of their earlier ancestors and a relatively sophisticated set of stone tools. An improved tool arsenal gave considerable advantage in the yield of food as well as the opportunity to spread to new habitats.

A major coevolution of culture and perhaps brain biology probably helps explain the origin of modern humans and their radiation. The human brain had already reached its present size before the second expansion out of Africa, with the first appearance of *H. sapiens*. During this period, modern human language likely evolved to its present degree of sophistication and was the major support of the second expansion. Although this is a speculation, there is hard evidence that in the period between 100,000 and 60,000 years ago the tool arsenal evolved substantially (*3*) and may have facilitated the expansion of modern humans. 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 (*9*). Use of rafts and boats must have been common because around 55,000 or 60,000 years ago *H.s.s.* passed several sea tracts to reach Australia (*10*). The development of navigational skills may have helped Africans to reach Southeast Asia (and finally Australia) along the southern coasts of Asia. A fast and precise system of verbal communication must have provided a great advantage as *H.s.s.* expanded to new territories and environmental niches, some of which had climate, flora, and fauna that were different from their original ones.

---

L. L. Cavalli-Sforza is at the Genetics Department, Stanford University, Stanford, CA 94305. P. Menozzi is at the Istituto di Ecologia, Università di Parma, Parma, Italy. A. Piazza is at the Dipartimento di Genetica, Biologia e Chimica Medica, Centro CNR per l'Immunogenetica e l'Istocompatibilità of the Università di Torino, Torino, Italy.

## Synthetic Genetic Maps of the World

Although the world was already inhabited by the end of the Paleolithic, the total population numbers were still low (*11*), perhaps between 1 and 10 million. Population movements after the Paleolithic did not have time to erase completely the genetic differences that were generated in the earlier period, as shown by simulations (*12, 13*), but added new genetic patterns that are still recognizable today. No single gene provides a clear record of the changes brought about by later migrations on the original differences, but a combination of the information from a large number of genes allows such reconstructions. We introduced the methodology of "synthetic" genetic maps, which are geographic maps of isopleths (lines of equal value) of principal component values (PCs), calculated as optimized linear functions of all available gene frequencies (*13*). The isopleth map that each PC generates is independent of all other maps and explains a calculable fraction of the total genetic variation. Each map is ranked in relative importance by the fraction of variance that is defined by its corresponding PC. We have shown by simulation that different PCs can separate expansions started at different times and places (*14*). In the simplest case, the rank of 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. It also depends on the demographic modalities in which the migration takes place, especially the rate of exchange with previous settlers. Another important factor is the increase in population density that can be attained when immigrants introduce a new culture and it reaches saturation. The rank of a PC is therefore associated with, but not exclusively dependent on, the antiquity of an expansion. For a successful analysis, the use of a large number of genes is essential. With the numbers we used, tests of statistical robustness (*15*) showed that maps of the highest PCs are highly reproducible.

The synthetic map that was obtained with the first PC depicts the major component in the genetic geography of the area that was studied. The world map (Fig. 1A) supports the hypothesis of an expansion of modern humans from Africa: One extreme PC value is in Africa and the other is in Australia. Taken in isolation, this synthetic map does not reveal which of these two poles was the origin of the expansion, but the archeological records show that Africa was occupied by modern humans (and their ancestors) long before Australia. The proportion of global genetic variation (observed with 120 genes) that was associated with the first PC is 34.8% (*16*). For such a large fraction of human variation, the west-to-east direction must have been the axis of a major expansion, the consequences of which were not entirely resolved by later migrations and local exchanges. In addition, adaptation to different climates cannot be the major determinant of variation

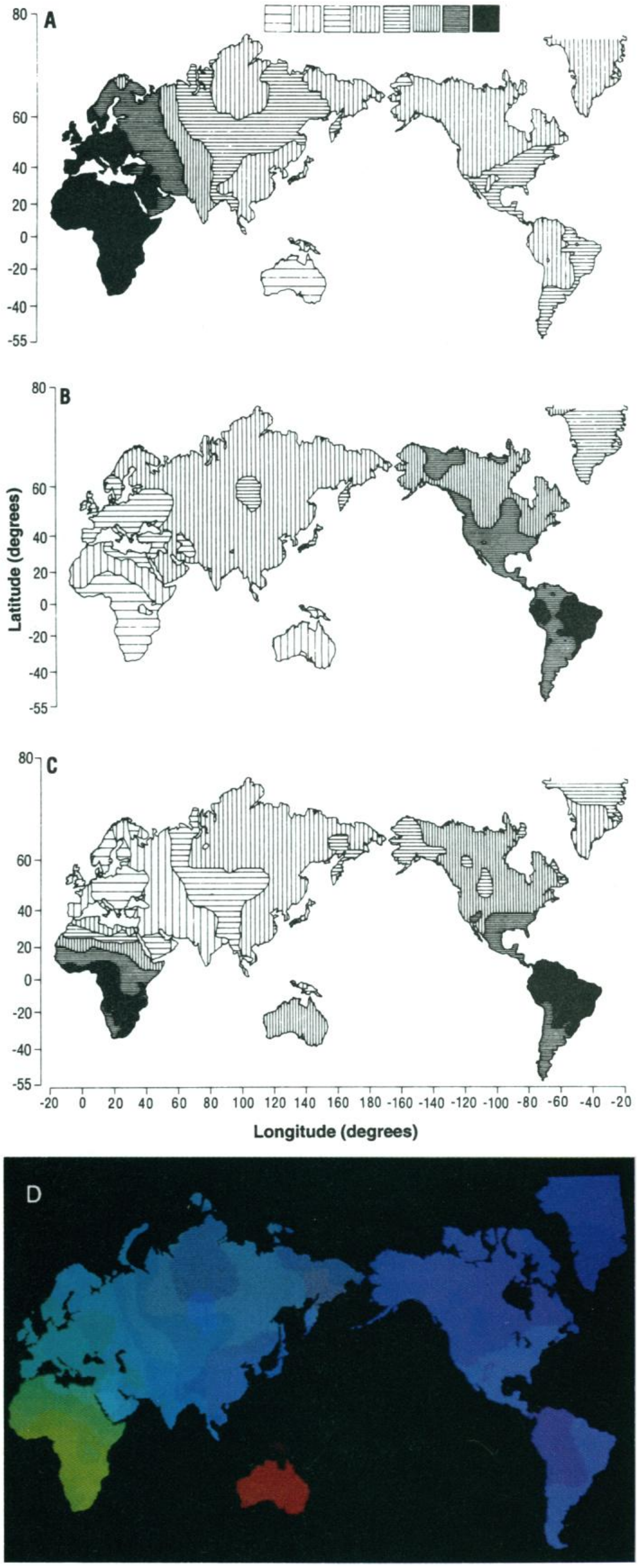

**Fig. 1.** Synthetic genetic maps of the world that represent 65% of the total genetic variation. (**A** to **C**) The interpolated values of the first to third PCs, which represent 34.8, 17.8, and 12.4% of the total genetic variation, respectively. The legend represents the PC scores, ranked 1 to 8 on an arbitrary scale. (**D**) The result of the image-processed superimposition of the previous maps, with green for the first, blue for the second, and red for the third PCs. This map conveys their global genetic information.

that is observed at the genetic level. This finding contrasts with observations on skull variation, which show considerable similarity between Australians and Africans and in which the first PC of both size and shape of the skull are highly correlated with climate (*17*). Skull morphology is, with skin color, a good indicator of past climate.

Lower PCs define patterns that are of progressively lesser importance to the explanation of the global genetic variation but that are statistically independent of the first ones. The second PC (Fig. 1B) shows a major difference between Africa and America, the last continent to be settled, and is responsible for 17.8% of the global variation. The third PC of the world (Fig. 1C) and the second PCs of Europe (Fig. 2B) and of Asia show that gene frequencies may have some association with climate. However, as indicated also in (*18*), anthropometrics seem more subject to natural selection for climate than are the genetic polymorphisms represented in our PC maps. Anthropometrics are therefore poor indicators of population history, which is best studied on selectively neutral markers, if possible. Figure 1D superimposes the first three principal components, with different primary colors for each. It accounts for 64% of the gene frequency variation.

## Expansions in Synthetic Maps

The PCs of partial regions of the world are easier to study than those of the whole world because they summarize a less complicated history. At this point some criteria for the interpretation of synthetic maps may be useful. Gradients of PC values that are formed by concentric rings around an area likely indicate an expansion whose origin is at the center of the rings. The center can have the highest or lowest PC value (represented as dark or light in the maps); the sign of PC values is assigned arbitrarily, and hence dark and light could be exchanged. A maximum or minimum (the darkest or lightest spot) indicates that the gene frequencies of the local population differ most from the average gene frequencies of the investigated region. A local maximum or minimum indicates a difference from neighboring populations. Every PC emphasizes the information generated by a specific set of genes in the specific region under study.

An expansion from a nuclear area will usually give rise to a genetic gradient with concentric rings that are visible in a synthetic PC map if several conditions are fulfilled: (i) Geographic isolation of the population at the center of the origin has previously determined an adequate difference in several of its gene frequencies from those of the neighboring areas. It is reasonable to assume that the origin of divergence is genetic isolation before the expansion, long and intense enough that the population gene frequencies drifted away from those of neighboring populations for a long time. Divergence might also be due to strong local natural selection, but a substantial number of genes must have been affected for a PC of high rank to be involved. Drift is more likely to cause local genetic divergence because it affects with equal probability all gene frequencies and PC values summarize variation of all genes, whereas selection that is caused by a specific environment is more likely to change only a few genes. (ii) Some cultural or biological development at the origin of the expansion has caused intensive population growth in this area. (iii) As soon as local saturation of population was reached at the center, the population expanded geographically to areas where the means of transportation and environmental conditions allowed migration. The repetition of cycles of growth and migration over many generations could carry the expansion far from the origin. (iv) The gradient has a short range if admixture with earlier settlers of the neighboring areas is high. The same result occurs if the cultural innovation responsible for population growth is easily passed culturally to neighbors. However, the gradient around the origin may be flat if the first condition is not met and if the earlier inhabitants of the areas that were settled by the expanding population were rare, did not mix with it, or were largely eliminated by it. (v) Later or independent migrations did not diminish, blur, or alter the gradient (*12, 14*).

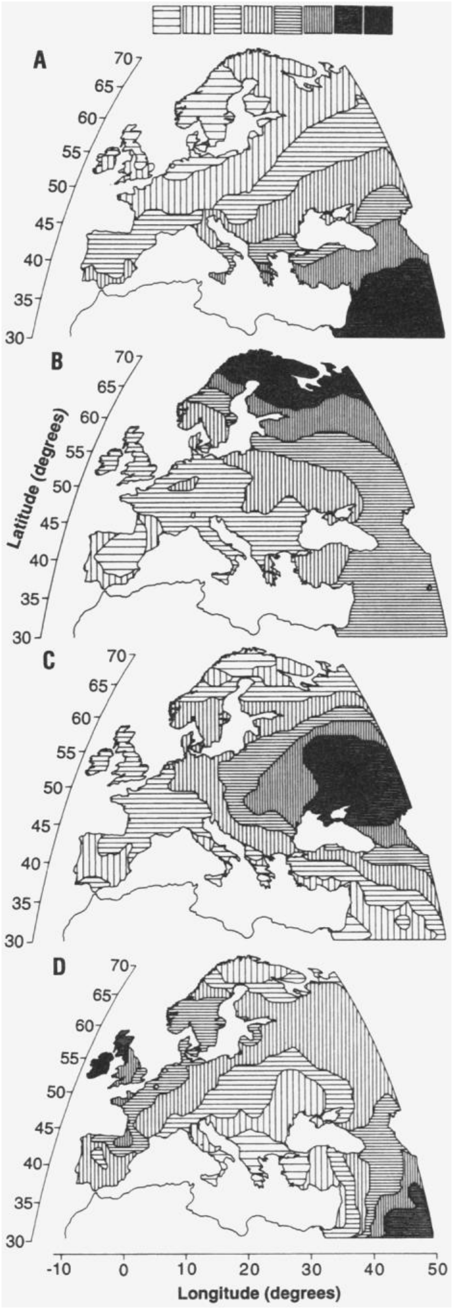

**Fig. 2.** Synthetic genetic maps of Europe. (**A** to **D**) The interpolated values of the first to fourth PCs, which represent 28.1, 22.2, 10.6, and 7.0% of the total genetic variation, respectively. Legend as in Fig. 1.

This may seem a rather restrictive list of conditions, but the first was normal in the Paleolithic when population densities were low; major innovations (the second condition) were not frequent, and the same is true of major expansions; and the third condition is an almost automatic consequence of the second. With regard to the fourth condition, we can study only observed gradients and cannot determine exactly how many we missed. The fifth condition is addressed by simulations (*14*) that have shown that gradients that are determined by important expansions are rather stable in time and that blurring due to independent migrations can be partially removed by use of PCs, which can separate one expansion from the other if their geographic origins are sufficiently distinct.

A maximum or a minimum in a synthetic map is not necessarily the result of a centrifugal event as described above. The cause could also be a centripetal population movement, as in the case of migration to capital cities or other centers that attracted many immigrants. Historical knowledge can in most cases confirm the origin of such spots. Examples of centripetal movement can be found in published synthetic maps of Europe for Paris and London (*13*) and for Rome and Naples (*19*). Finally, there can be PC gradients that are not determined by expansions but by, for example, special environmental parameters such as latitude.

## Neolithic and Post-Neolithic Expansions

The change from food collection (hunting and gathering) to food production (agriculture and animal breeding) is called in Europe the Neolithic transition and began in a nuclear area located in the Middle East around 10,000 years ago. Cereals and domesticated animals from this area spread slowly in all directions soon after that time. The best known expansion, that toward Europe, took place at the average rate of 1 km per year (*20*). The question was asked whether the spread from the area of origin was cultural (that is, was it farming that diffused) or demic (was it farmers who radiated) (*21*). The rate of the diffusion of agriculture is quantitatively compatible with a demic spread (*20*).

We originally developed synthetic maps (*13*) as a device to test further this hypothesis. It was suggested that if farmers spread from the Middle East in the Early Neolithic to all of Europe, it might still be possible to detect 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 (*20*), and the genetic maps. The correlation between the archeological map and the synthetic map that was obtained with the first PC was remarkably clear (*13*). On this evidence it was concluded that the diffusion of farming from the Middle East to Europe was a spread of farmers and not just technology. This result was confirmed by Sokal and Menozzi with spatial autocorrelation methods (*22*). Sokal, Oden, and Wilson (*23*) used 59 markers, more than the 39 we used originally (*13*), and found that archeological time distances and genetic distances, partialling out geographic distances, were highly correlated ($P < 10^{-5}$). A new map of the first PC of Europe (Fig. 2A) (*16*) based on 95 genes instead of the 39 we first used is in excellent agreement with the archeological information. The first PC has a high and significant correlation (*15*) with the dates of first arrival of Neolithic farming [given in (*20*)].

Maps of Europe that were obtained with PCs of lower rank are also of great interest. The second and third synthetic maps (*16*) confirm and extend earlier maps (*13*) that had less than half the statistical information. The second map of Europe (Fig. 2B) shows a gradient from north to south, which may reflect a climatic effect but may also be attributable at least in part to ethnic origin. In the north of Europe live Lapps, who have genetic similarity with other northern Eurasian populations from Siberia and speak a language from the Uralic family, found today mostly east and west of the Urals. Lapps have undergone substantial admixture with Scandinavians (*24*) and, in spite of their probable west Siberian origin, are now classified as Europeans (*16, 25*). However, as the extreme European outliers, the Lapps diverge considerably from all peoples of that region. Southern Scandinavians expanded into the area occupied in the north by people of Uralic origin like the Lapps, and this expansion established a local genetic gradient. Other people of Uralic origin are located farther east, on both sides of the Ural chain, and resemble North Asians located even further east. Thus, the north-south gradient may be a result of selective effects of climate on North Asian populations like those that speak Uralic languages and that admixed with the North Europeans who came later to the area.

The third PC of Europe (Fig. 2C) strongly suggests another expansion, which was visible in our earlier analysis (*13*). The improved precision of more than twice the number of genes led to a shift in the origin about 500 km in the southeast direction in the region of the Don River. This is the same general area where archeological data place the beginnings of the "Kurgan" culture, believed to have started around 6000 to 4000 years ago by pastoral nomads of the Eurasian Steppes. Kurgan is the Russian name for a burial mound, common in the central Russian Steppes. A connection was suggested between the Kurgan culture and migrations of speakers of Indo-European languages by Gimbutas (*26*).

Recently, Renfrew suggested that Indo-European languages came to Europe from Anatolia with the spread of farmers (*27*). His rationale was that farmers did spread according to the demic hypothesis, as discussed above, and under such circumstances they must have carried to Europe their original language. This language must have undergone considerable evolution in the long process of spread and replaced almost all languages spoken in Europe before the Indo-European language, with one clear exception: Basque. Some linguists have independently argued for an Anatolian origin of Indo-European languages (*28*).

Hypotheses on the place of origin of a language are notoriously difficult to support or disprove. The extraordinary variety of homelands suggested for Indo-Europeans (*29*) emphasizes the uncertainty of such choices. The problem is compounded by the possibility, which has perhaps not been sufficiently recognized, 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 4000 and 2500 B.C.), which seems to have greatest support from archeological and other considerations discussed in (*29*), is really not an alternative to the Anatolian origin suggested by Renfrew and others; both hypotheses can be correct. The Kurgan origin would have been later and secondary to the spread of agriculture from the Middle East to the steppes and the establishment there of the new economy of pastoral nomadism.

The success of pastoral nomads from central Eurasia was not limited to the postulated expansion of the Kurgan culture to Europe but included a great number of other, some quite recent, radiations. These expansions reflect an adaptation of farmers to prairies, savannas, or even arid regions that were not favorable to farming but lent themselves to the breeding of useful indigenous animals. Unquestionably, however, one of the most significant secondary developments of agriculture was pastoral nomadism in the Eurasian Steppes, which extend from Romania to Manchuria almost uninterrupted. The domestication of the horse took place in the Western Steppes about 5000 years ago or earlier and was soon followed by the use of horses in the military (*30, 31*). The migrations in large numbers of pastoral nomads demanded the development of considerable mobility, logistics, and military skills. War chariots and later cavalry were important innovations that gave considerable military and political power to nomadic pastors of central Asia. The Western Steppes were originally settled by Caucasoid people who spoke Indo-European languages and who spread in several directions, including east (toward the Altai) and south (to Iran and India). After about 300 B.C., nomads from the Eastern Steppes, Mongoloid people who spoke Altaic languages, also began major expansions and in the process became the scourge of Asia and Europe.

A symbiosis developed between pastoral nomads and farmers who originally owned the land (*31*). Nomads often generated a hierarchical society in which their aristocracy formed the highest caste, and the farmers were organized in various lower classes 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 followers of the Hindu religion never acquired control or lost it to the Moslems.

The expansions of pastoral nomads from the Eurasian Steppes may well explain the migrations of people speaking Indo-European languages not only to India but also to Europe, as suggested by Gimbutas (*26*) but rejected by other archeologists including Renfrew (*27*). The synthetic maps based on the third (Fig. 2C) and sixth PCs of Europe indicate expansions that correspond geographically to two of the three origins of Kurgan expansions suggested by Gimbutas. On the other hand, these may be partially confused with the later expansions of, for example, the Scythians and of barbarians who infiltrated or conquered the Roman

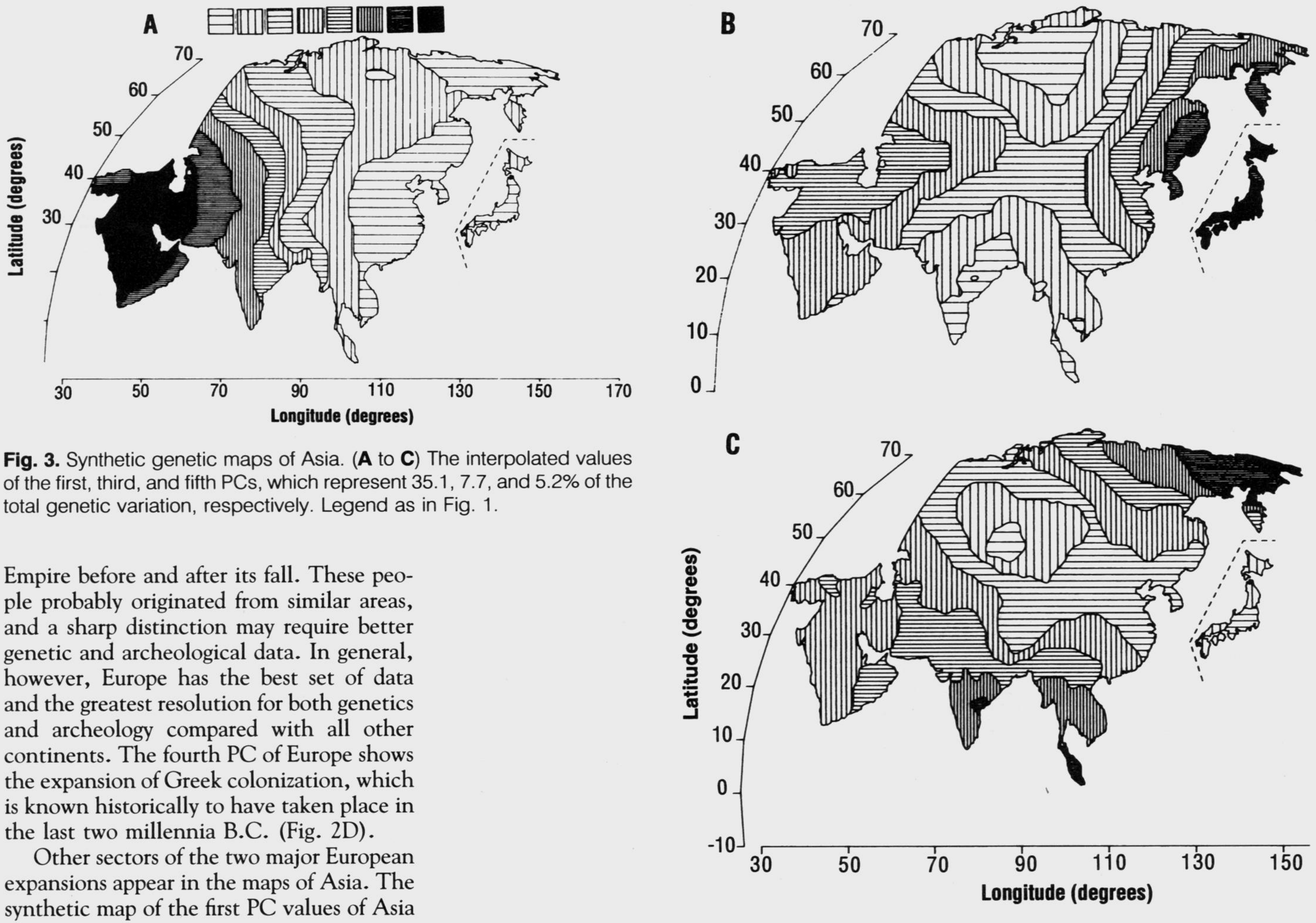

**Fig. 3.** Synthetic genetic maps of Asia. (**A** to **C**) The interpolated values of the first, third, and fifth PCs, which represent 35.1, 7.7, and 5.2% of the total genetic variation, respectively. Legend as in Fig. 1.

Empire before and after its fall. These people probably originated from similar areas, and a sharp distinction may require better genetic and archeological data. In general, however, Europe has the best set of data and the greatest resolution for both genetics and archeology compared with all other continents. The fourth PC of Europe shows the expansion of Greek colonization, which is known historically to have taken place in the last two millennia B.C. (Fig. 2D).

Other sectors of the two major European expansions appear in the maps of Asia. The synthetic map of the first PC values of Asia (Fig. 3A) shows a strong gradient from west to east that has an apparent origin in the Middle East. This likely represents at least part of the eastern sector of the expansion of agriculture from the Fertile Crescent, which certainly reached as far as India in the east and the steppes in the north. The center of agriculture in the Middle East, which we have seen spread toward Europe, extended also in the east to Iran, Afghanistan, Pakistan, and India. It has been suggested that Dravidian languages were spread by farmers of the eastern part of the Fertile Crescent (*32*) and were later replaced by Indo-European languages that were introduced by pastoral nomads, except in south India and a few northern areas. The map of the Asian fifth PC component (Fig. 3C) shows a possible nuclear area in a region near that of the Afanasievo culture, which was probably also an important center of expansion of steppe nomads.

The synthetic map of the second Asian PC values is an almost perfect gradient north to south and is not dissimilar from the map of the second world PC. It may in part represent a selection by climate but also shows a gradient of expansion toward the north from the central part of Asia, which was certainly populated before migration to Siberia took place. In the map of the third PC (Fig. 3B) there is an apparent center of expansion in middle Japan. In the last glacial period there was an almost complete continuity between Japan and continental Asia in both the north and the south, and the Sea of Japan was practically an internal sea. From an archeological survey of almost 28,000 sites (*33*), the population growth since 10,000 years ago in Japan has been reconstructed at least in the regions that are still above water. The survey shows a stable population growth at about 1% per year from an estimated number of Paleolithic inhabitants of around 3000. About 4000 years ago the Japanese population reached its maximum of about 260,000, with a density of 0.89 per square kilometer and considerable variation from region to region. This period is called the Neolithic, although there was no agriculture until about 1700 years ago. Foraging and fishing provided for large population densities in Japan without the use of agriculture. A similar development took place on the Pacific Coast of northwest North America. Thus, there is a strong indication of an expansion west and north from Japan. The date of this movement is difficult to establish, but the time of the population peak of about 4000 years ago may be a reasonable estimate for the expansion. An earlier spread during the Late Paleolithic, however, cannot be excluded.

The genetic map of Africa is reasonably well known except in the Sahara, which is currently inhabited only in restricted zones, some of which are almost unstudied. The synthetic map of the first PC values shows the distinction between the Caucasoid populations that inhabited the northern part of the Sahara and black sub-Saharan populations (Fig. 4A). The Sahara was not a desert until a few thousand years ago and was the seat of agricultural developments, especially cattle breeding (*16*). With the desertification of the Sahara 3000 to 4000 years ago, cattle had to migrate south and farmers (mostly dark-skinned) started domesticating plants such as sorghum and millet in the Sahel region, immediately south of the Sahara. Many of these new agricultural developments were in West Africa, where a well-known expansion to the east and south began. This radiation from West Africa is visible on the second synthetic map (Fig. 4B). This map also shows that the southwestern part of South

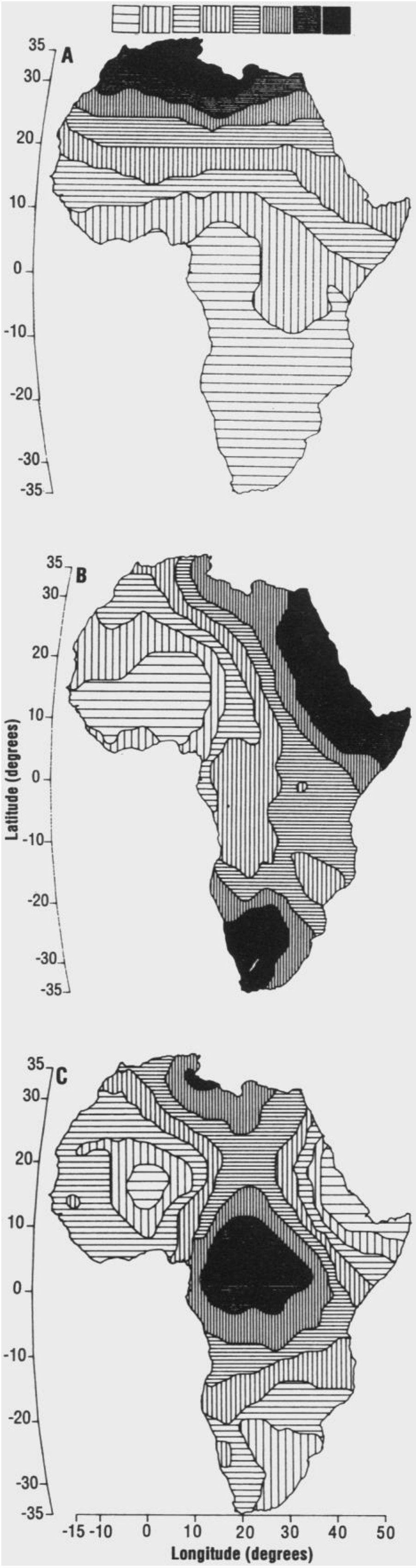

**Fig. 4.** Synthetic genetic maps of Africa. (**A** to **C**) The interpolated values of the first, second, and fourth PCs, which represent 34.6, 18.6, and 7% of the total genetic variation, respectively. Legend as in Fig. 1.

Africa is quite different from West Africa. It is occupied today mostly by speakers of Khoisan languages (Bushmen and Hottentots), who show some remote genetic similarity to East Africans (*16*) but have a different linguistic background. Another diffusion of farmers (visible in Fig. 4C) started at least 3000 years ago and is called the Bantu expansion because it originated in Cameroon near the boundary with Nigeria by people who spoke Bantu or Bantu-like languages and spread them across central and most of South Africa (*34*).

Another interesting anomaly is observed in the fourth synthetic map of Africa (Fig. 4C), which reveals a possible center of expansion in a region shared by Burkina Faso, Niger, and Mali. This area has other genetic peculiarities: the high frequencies of two hemoglobin genes, C and D (not used for calculation of the PC maps), whose origins are centered near the West African anomaly of the fourth PC map. Hemoglobin C is centered slightly below, and hemoglobin D slightly above, the center of the PC anomaly. These two hemoglobins are perhaps unrelated to malarial resistance (*35*). Genetic drift, perhaps in conjunction with selection, may have caused hemoglobins C and D to reach high frequencies in the populations that were later involved in the agricultural expansion that is indicated (Fig. 4C). This expansion later spread them passively over a wider area.

The synthetic maps of the Americas are more complicated, mostly because of extreme drift (*16*, *36*). The first synthetic map (Fig. 5A) shows the genetic differences among the three major groups of American natives: Eskimos in the extreme north, Na-Dene below them, and American Indians elsewhere. This classification was suggested by the linguistic work of Greenberg, odontological work of Turner, and preliminary genetic work of Zegura (*37*). We have fully confirmed this division on the basis of genetic data (*16*). The second synthetic map (Fig. 5B) indicates heavy hybridization with European settlers on the East Coast of North America and the basic homogeneity of the Andean region compared with the rest of South America. Analysis of the whole continent, however, fails to isolate regions of agricultural spread, whose origins on the Mexican plateau and the northern Andes are well known. Agriculture spread late from Mexico to North America, perhaps because it was difficult to cross the Sonoran Desert in the north of Mexico and because the regions that bordered Mexico farther north were also arid. It was only about 5000 years after beginning of agriculture in Mexico that it began to develop in the southern United States and to spread across the prairies (*38*). The diffusion from the Andes to the Amazonian region and the Caribbean is still not completely understood [but see (*39*)]. Oceania and the Pacific area are more difficult to study because the genetic data are mostly poor or irregularly distributed and reflect several expansions (*16*).

## Major Expansions and Gene-Culture Coevolution

The genetic map of the world seems to trace many human expansions. Each was most probably the result of some technological innovation or innovations that generated population growth and helped people spread from their nuclear origin to new regions. Some of these new areas may have been unoccupied or only sparsely settled. Especially in the latter case, genetic clines must have formed. We can still detect them

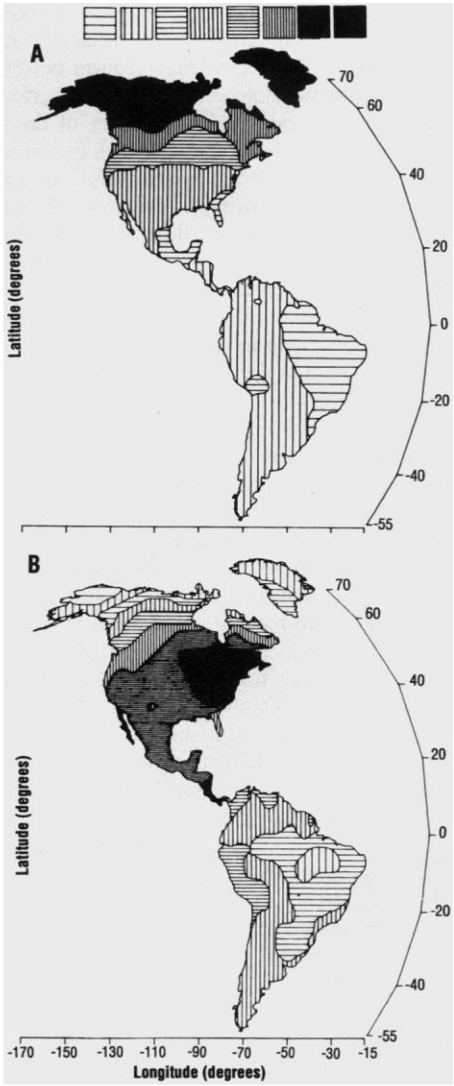

**Fig. 5.** Synthetic genetic maps of America. (**A** and **B**) The interpolated values of the first and second PCs, which represent 32.6 and 12.7% of the total genetic variation, respectively. Legend as in Fig. 1.

because it takes a long time for local short-distance migration to erase clines that extend over wide areas (14) and because synthetic map analysis allows us to unravel the independent, superimposed waves of migration.

The traces of some major expansions that occurred in the Paleolithic highlight the major differences we still see among aboriginal populations of entire continents and parts of them. After the Paleolithic, the major expansions were determined by the development of agriculture, which arose at similar times in widely different regions. The earliest of these agricultural communities were in the Middle East, China (several centers, one in the north and at least two in the south), and Mexico and the northern Andes. Spread from the Middle East is clearly visible toward Europe, Central and South Asia, and to some extent North Africa. Here they probably overlap with Arab and Bedouin migrations to North Africa in the 7th and 11th centuries A.D. (16).

Genetic data are not yet adequate to clarify how they were affected by the Chinese agricultural expansions. North and South China have different ecologies, neolithic histories, and genetic and cultural geographies. At least initially, the peoples in this region grew different crops (millet in the north and rice in the south). Current genetic data show a sharp difference between North and South Chinese (16, 40). Higher resolution of genetic and archeological maps that is not possible with present East and Southeast Asian data is necessary for a better analysis.

Some of these and other major expansions (Table 1) were made possible by improvement in transportation rather than in food collection or production. Some may have been due to other types of technological improvements, such as house building and the production of clothing. Some expansions occurred before or shortly after the beginning of historical records, and others later were triggered by military developments. We suspect that there were several prehistorical expansions other than those listed in Table 1 that cannot yet be confirmed or dated for lack of relevant archeological information. Genetic maps can provide indications that are useful for the detection of previously unsuspected expansions. They can also help to locate the centers of origin of demic expansions that were detected in other ways.

Improved knowledge of genetic geography is essential to sharpen this method of study. For the best analysis of a particular area, study should be restricted to a region around that of interest so as to minimize interference from other expansions. In our effort we have concentrated on aboriginal peoples, who were in place before the major geographic discoveries of the 15th and 16th centuries A.D. Therefore, they do not include (or do so only indirectly) the expansions of colonial powers to other continents, in particular the Americas and Australia. These migrations are reasonably well known historically and have had a major genetic impact on modern populations. Study of these expansions could benefit from historical demographic information, which is sometimes available.

**Table 1.** Major prehistoric expansions.

| Center of origin | Area of expansion | Time* | Technologies |
|---|---|---|---|
| Africa (*H. erectus*) | Old World | >1 Ma | Stone tools |
| Africa (*H.s.s.*) | Entire world | 100 to 30 ka | New tools, more advanced language, and navigation |
| Middle East | Europe, North Africa, and Southwest Asia | 10,000 to 5,000 ya | Farming and domestication (wheat, barley, goats, sheep, and cattle) |
| North China | North China | 9,000 to 2,000 ya | Farming and domestication (millet and pigs) |
| South China | Southeast Asia | 8,000 to 3,000 ya | Farming and domestication (rice, pigs, and buffalo) |
| Central America and North Andes | Americas | 9,000 to 2,000 ya | Farming (corn, squash, and beans) |
| West Africa† | Sub-Saharan Africa | 4,000 to 300 ya | Farming (millet, sorghum, cowpea, and gourd) |
| Eurasian Steppes | Eurasia | 5,000 to 300 ya | Pastoral nomadism (horses and warfare) |
| Southeast Asia or Philippines | Polynesia | 5,000 to 1,000 ya | Oceanic navigation |
| Greek colonization | Mediterranean | 4,000 to 2,400 ya | Navigation and trade |

*Abbreviations: Ma, million years ago; ka, thousand years ago; ya, years ago. †Includes the Bantu expansion.

A simple mathematical theory that was developed by Fisher for the spread of advantageous genes (41) is applicable to population expansions (42). This theory can help predict the average rate of demic expansions, as was done for the diffusion of agriculture to Europe (20): the radial rate of diffusion is the geometric mean of the growth rate and the migration rate. This radial rate is the initial rate of increase of a logistic growth rate and can be higher than rates that are usually measured archeologically. It could be close to 3% per year, as observed by Birdsell (43). Migration rates are more difficult to measure (20).

Demic diffusion entails the spread of a population's culture, including its language. The hypothesis by Renfrew (27) that Neolithic farmers who radiated from the Middle East spread their languages is very likely correct. However, it is difficult to test whether the languages we still speak descend from theirs because there is, by definition, no written record of prehistoric events, and prehistoric languages are inferred indirectly. Moreover, because of their development over long periods such as that of the spread of Neolithic farming to Europe (from 9500 to 5500 years ago), the initial languages may have changed and differentiated beyond recognition. Languages were carried by European settlers in their spread after the great geographic discoveries of the Renaissance. A widely accepted prehistoric example that has been confirmed genetically is that of the diffusion of Bantu languages by slash and burn African farmers (16). Other cultural traits, especially highly conserved ones related to family structure and kinship, have diffused with African farmers (44). Especially in traditional societies, vertical cultural transmission (from parents to children) is common. Horizontal transmission (between unrelated people) is rare in economically primitive societies but is common in modern ones (45). Cultural traits transmitted vertically, and therefore in a way more similar to that of genes, are likely to share the same evolutionary history. Under these conditions one can expect coevolution of genes and culture. Languages will also follow this coevolutionary pattern (7, 46), except for episodes of horizontal transmission such as whole language replacement.

## REFERENCES AND NOTES

1. N. Eldredge and S. J. Gould, in *Models in Paleobiology*, T. J. Schopf, Ed. (Freeman, San Francisco, 1972), pp. 82–115.
2. G. Brauer, in *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*, P. Mellars and C. Stringer, Eds. (Princeton Univ. Press, Princeton, NJ, 1989), pp. 123–154; P. G. Rightmire, *ibid.*, pp. 109–122; C. B. Stringer, *ibid.*, pp. 232–244.
3. R. G. Klein, *ibid.*, pp. 529–546.
4. F. Weidenreich and C. S. Coon, *The Living Races of Man* (Knopf, New York, 1965); M. H. Wolpoff, in (*2*), pp. 62–108.
5. M. Barinaga, *Science* **255**, 686 (1992); A. R. Templeton, *ibid.*, p. 737; S. B. Hedges, S. Kumar, K. Tamura, M. Stoneking, *ibid.*, p. 737.
6. R. L. Cann, M. Stoneking, A. C. Wilson, *Nature* **325**, 31 (1987); L. Vigilant, M. Stoneking, H. Harpending, K. Hawkes, A. C. Wilson, *Science* **253**, 1503 (1991).
7. L. L. Cavalli-Sforza, A. Piazza, P. Menozzi, J. Mountain, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6002 (1988).
8. A. M. Bowcock *et al.*, *ibid.* **88**, 839 (1991); J. L. Mountain, A. A. Lin, A. M. Bowcock, L. L. Cavalli-Sforza, *Philos. Trans. R. Soc. London* **337**, 159 (1992).
9. G. L. Isaac, in *Origins and Evolution of Language and Speech*, S. R. Arnald, H. D. Steklis, J. Lancaster, Eds. (New York Academy of Sciences, New York, 1976), vol. 280, pp. 275–288.
10. J. B. Birdsell, in *Sunda and Sahul: Prehistoric Studies in Southeast Asia, Melanesia and Australia*, J. Allen, J. Golson, R. Jones, Eds. (Academic Press, London, 1977), pp. 113–167; R. M. Jones, in (*2*), pp. 743–782.
11. J. N. Biraben, *Population* **4**, 1 (1980); F. A. Hassan, *Demographic Archeology* (Academic Press, New York, 1981).
12. L. Sgaramella-Zonta and L. L. Cavalli-Sforza, in *Genetic Structure of Populations*, N. E. Morton, Ed. (Univ. of Hawaii Press, Honolulu, 1973), pp. 128–135.
13. P. Menozzi, A. Piazza, L. Cavalli-Sforza, *Science* **201**, 786 (1978).
14. S. Rendine, A. Piazza, L. L. Cavalli-Sforza, *Am. Nat.* **128**, 681 (1986).
15. A. Piazza, S. Rendine, P. Menozzi, J. Mountain, L. L. Cavalli-Sforza, in preparation.
16. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, in press).
17. W. W. Howells, *Pap. Peabody Mus. Archaeol. Ethnol. Harv. Univ.* **67**, 1 (1973); *ibid.* **79**, 1 (1989); C. R. Guglielmino Matessi, P. Gluckman, L. L. Cavalli-Sforza, *Am. J. Phys. Anthropol.* **50**, 549 (1979).
18. A. Piazza, P. Menozzi, L. L. Cavalli-Sforza, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2638 (1981).
19. A. Piazza, N. Cappello, E. Olivetti, S. Rendine, *Ann. Hum. Genet.* **52**, 203 (1988).
20. A. J. Ammerman and L. L. Cavalli-Sforza, *Neolithic Transition and the Genetics of Populations in Europe* (Princeton Univ. Press, Princeton, NJ, 1984).
21. ———, in *The Explanation of Culture Change*, C. Renfrew, Ed. (Duckworth, London, 1973), pp. 343–357.
22. R. R. Sokal and P. Menozzi, *Am. Nat.* **119**, 1 (1982).
23. R. R. Sokal, N. L. Oden, C. Wilson, *Nature* **351**, 143 (1991); *ibid.* **355**, 214 (1992).
24. C. R. Guglielmino Matessi, A. Piazza, P. Menozzi, L. L. Cavalli-Sforza, *Am. J. Phys. Anthropol.* **83**, 57 (1990).
25. L. L. Cavalli-Sforza and A. Piazza, *Eur. J. Hum. Genet.* **1**, 3 (1993).
26. M. Gimbutas, *The Civilization of the Goddess: The World of Old Europe* (Harper, San Francisco, 1991), chap. 10.
27. C. Renfrew, *Archaeology and Language: The Puzzle of Indo-European Origins* (Jonathan Cape, London, 1987).
28. A. B. Dolgopolvsky, in *Mediterr. Lang. Rev.* **3**, 7 (1988); M. Kaiser and V. Shevoroshkin, *Annu. Rev. Anthropol.* **17**, 309 (1988); T. V. Gamkrelidze and V. V. Ivanov, *Sci. Am.* **262**, 110 (March 1990).
29. J. P. Mallory, *In Search of the Indo-Europeans: Language, Archaeology and Myth* (Thames and Hudson, London, 1989).
30. M. Zvelebil, in *The Cambridge Encyclopedia of Archaeology*, A. Sherrat, Ed. (Crown, New York, 1980), chap. 38; J. M. Diamond, *Nature* **350**, 275 (1991).
31. A. M. Khazanov, *Nomads and the Outside World* (Cambridge Univ. Press, New York, 1984).
32. L. L. Cavalli-Sforza, *Munibe* (*Suppl.*) **6**, 129 (1988); C. Renfrew, *Sci. Am.* **261**, 106 (October 1989).
33. S. Koyama, *Senri Ethnol. Stud.* **2**, 1 (1978).
34. J. H. Greenberg, *The Languages of Africa* (Univ. of Indiana Press, Bloomington, 1963).
35. F. B. Livingstone, *Abnormal Hemoglobins in Human Populations* (Aldine, Chicago, 1967); L. L. Cavalli-Sforza and W. F. Bodmer, *The Genetics of Human Populations* (Freeman, San Francisco, 1971).
36. D. H. O'Rourke, B. K. Suarez, 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 Univ. Press, Oxford, 1988).
37. J. H. Greenberg, G. C. Turner II, S. L. Zegura, *Curr. Anthropol.* **27**, 477 (1986).
38. T. P. Culbert, in *Ancient South Americans*, J. D. Jennings, Ed. (Freeman, New York, 1983), chap. 2.
39. D. W. Lathrap, in *Origins of Agriculture*, C. A. Reed, Ed. (Mouton, The Hague, Netherlands, 1977), pp. 713–751.
40. N. Saitou, K. Tokunaga, K. Omoto, in *Isolation and Migration*, R. F. Roberts, Ed. (Cambridge Univ. Press, Cambridge, 1991).
41. R. A. Fisher, *Ann. Eugen. (London)* **7**, 355 (1937).
42. J. Skellam, in *The Mathematical Theory and Dynamics of Biological Populations*, M. Bartlett and R. Hiorns, Eds. (Academic Press, London, 1973).
43. J. P. Birdsell, *Cold Spring Harbor Symp. Quant. Biol.* **22**, 47 (1957).
44. C. R. Guglielmino Matessi, C. Viganotti, L. L. Cavalli-Sforza, "Spatial Distributions and Correlations of Cultural Traits in Africa" (Monogr. 383, Consiglio Nazionale delle Ricerche, Pavia, Italy, 1983).
45. L. L. Cavalli-Sforza and M. Feldman, *Cultural Transmission and Evolution: a Quantitative Approach* (Princeton Univ. Press, Princeton, NJ, 1981); B. S. Hewlett and L. L. Cavalli-Sforza, *Anthropology* **88**, 922 (1986).
46. L. L. Cavalli-Sforza, *Sci. Am.* **265**, 104 (November 1991).

47. Supported by National Institutes of Health grant GM 20467 and 10452 and Ministero Università Ricera Scientifica Tecnica (Italy) and Consiglio Nazionale delle Ricerche projects "Biotechnology and Bioinstrumentation" and "Biological Archive." Computer time and graphics were granted by CSI-Piemonte (Torino, Italy).