

# THE NEOLITHIC INVASION OF EUROPE

---

**Martin Richards**

*Department of Chemical and Biological Sciences, University of Huddersfield,  
Queensgate, Huddersfield HD1 3DH, United Kingdom; email: m.b.richards@hud.ac.uk*

**Key Words** Neolithic farmers, Mesolithic foragers, mitochondrial DNA,  
Y chromosome, phylogeography

■ **Abstract** Who are Europeans? Both prehistoric archaeology and, subsequently, classical population genetics have attempted to trace the ancestry of modern Europeans back to the first appearance of agriculture in the continent; however, the question has remained controversial. Classical population geneticists attributed the major pattern in the European gene pool to the demographic impact of Neolithic farmers dispersing from the Near East, but archaeological research has failed to uncover substantial evidence for the population growth that is supposed to have driven this process. Recently, molecular approaches, using non-recombining genetic marker systems, have introduced a chronological dimension by both allowing the tracing of lineages back through time and dating using the molecular clock. Both mitochondrial DNA and Y-chromosome analyses have indicated a contribution of Neolithic Near Eastern lineages to the gene pool of modern Europeans of around a quarter or less. This suggests that dispersals bringing the Neolithic to Europe may have been demographically minor and that contact and assimilation had an important role.

## INTRODUCTION: FARMERS OUR ANCESTORS?

Prehistoric archaeology grew up under the shadow of nationalism, providing the means by which the newly established European nation-states could create a unitary past for their peoples (Trigger 1989, Kohl & Fawcett 1995). There was traditionally a tendency, therefore, for the narratives of European prehistorians to divide their actors into “us” and “them.” At the deepest level, as Zvebil (1995a) argues, this amounted to a founding myth for European culture and civilization that placed extraordinary emphasis on the Neolithic—a myth that idolizes farmers at the expense of hunting and foraging ways of life.

There are, Zvebil believes, three particular reasons why this may have happened. The first arose from the prejudice against hunter gatherers that emerged to justify the colonial persecution of hunter-gatherer communities in the New World (Gamble 1992). The second is the rise of urbanism and the resulting idealization of the rural way of life. The third is the need of many of the new nation-states, especially in central and northeast Europe, to forge a national identity among their predominantly peasant populace. The effects were felt not only in archaeology but

also in literature from Wordsworth and Hardy to Tolstoy, and much more widely throughout the popular culture.

In archaeology, of course, the triumph of these views came with V.G. Childe's proposal that the appearance of the first farming communities in Europe—the future Marxist Childe's "Neolithic revolution"—was the result of immigration of populations from the Near East (Childe 1925). He proposed that following the onset of the Neolithic in the Near East at the end of the last glaciation, about 10,000 years ago, farmers dispersed north and west into Europe, replacing the indigenous, Mesolithic hunting and foraging "savages" by virtue of their superior technology and culture. Zvelebil argues, along with a number of Palaeolithic and Mesolithic archaeologists, that the contribution of Europe's earlier, indigenous inhabitants to European society has been underestimated ever since.

This presumption in favor of what Zvelebil calls "farmers our ancestors" was perpetuated by both Neolithic archaeologists and population geneticists, as the latter became involved in prehistoric reconstruction from the 1970s onward. As recently as the 1980s, Bradley (1984) was able to make his famous observation that "in the literature as a whole, successful farmers have social relations with one another, while hunter-gatherers have ecological relations with hazelnuts." Yet the period since his remark has begun to witness a sea change in our view of the Neolithic and, finally, a rehabilitation of Europe's hunter-gatherer past. This has happened for two reasons: first, a sustained critique of the mass-migration perspective from the archaeological community, stressing in particular the importance of the social context of the Neolithic transition, and, second, by the arrival of molecular genetics on the scene.

The image of an invasion force marching across Europe at the beginning of the Neolithic has always been a caricature, but the question of the relative contribution of newcomers and natives to the European Neolithic has been vigorously fought over during the last 15 years. This question in turn has implications for explaining the social changes that took place with the shift to the Neolithic (e.g., Thomas 1996, 1998; Zvelebil 2000), and perhaps also for the arrival of new languages (Renfrew 1987, Zvelebil 1995b). On the genetic side, however, it appears that some consensus may finally be emerging. This article aims to review that part of the story, in order to assess the state of the argument over "farmers our ancestors" today.

## THE WAVE OF ADVANCE

The subject of the genetic history of Europe was more or less created by Luca Cavalli-Sforza and his colleagues in the 1970s. Their work has cast a long shadow, to the extent that today's genetic researchers into the ancestry of Europeans, and even many archaeologists, continue to discuss their subject matter in the terms that they laid down. Cavalli-Sforza's work was pioneering in two ways. It was the first sustained attempt to apply genetic data to a question of major archaeological interest. In addition, however, it helped to bridge the gap between genetics and archaeology by being carried out in collaboration with an archaeologist, Albert

Ammerman. The partnership began in 1970 and culminated in 1984 with *The Neolithic Transition and the Genetics of Populations in Europe*, which remains a seminal work in archaeogenetics.

Ammerman & Cavalli-Sforza (1984) drew upon the new processual archaeology of Lewis Binford (1968) for a scientific model explaining the origins and spread of farming in western Eurasia. Binford's density equilibrium model suggested that farming had arisen in the Near East in areas of optimal wild resources, leading to sedentism, local population growth, and expansion into more marginal environments. Ammerman and Cavalli-Sforza accepted the central role of sedentism, population growth, and the resulting resource pressure in the early farming communities.

Cavalli-Sforza had studied under the founder of orthodox statistics, R.A. Fisher, and was keen to apply quantitative methods to the study of prehistoric Europe. His work with Ammerman began by measuring the rate of the spread of farming into Europe, drawing on the newly available radiocarbon maps compiled by Clark (1965), showing an east-west trend in first Neolithic dates. They observed that there was a complex of elements at Neolithic sites in Europe—the so-called Neolithic package. This included cereal crops—especially emmer wheat, einkorn wheat, and barley, whose wild progenitors occurred only in the Fertile Crescent of the Near East—domestic animals, pottery, ground and polished stone tools, and houses. To maximize their data set, Ammerman and Cavalli-Sforza sometimes chose to rely upon single items, focusing on cereals where possible, as markers for the spread of the Neolithic. They developed the use of isochron maps, plotting similarly dated sites on a map of Europe as a series of isochrons or contour lines. The result was a remarkably uniform rate of about one kilometer per year, or 25 kilometers per generation, with minor variations in different parts of the continent. The entire process, from Greece to the British Isles, had taken place in about 2500 years. This is in fact rather rapid: less time, for example, than it took farming to spread throughout the Near East, where it had originated (Price 2000a).

This discovery that the rate of spread across Europe appeared to be roughly constant suggested a single overarching mechanism, a diffusionary process rather than an old-fashioned model of directed colonization. Ammerman and Cavalli-Sforza introduced the expression demic diffusion to contrast the immigration of farmers themselves with the spread of farming as an idea through the indigenous hunter-gatherer populations—cultural diffusion. Demic diffusion could imply a rather traditional model of migration and colonization, but Ammerman and Cavalli-Sforza argued that the observed rate of spread suggested something different: a “wave of advance.”

They took the wave of advance model from Cavalli-Sforza's mentor, Fisher, who had used it to describe the spread of an advantageous gene through a population. The wave of advance combines two features: logistic population growth and random local migratory diffusion or range expansion. The population growth was explained as the result of agriculturalist surpluses and storage in Neolithic societies, which allowed the carrying capacity of the land to rise. The outcome of

growth combined with range expansion is a radial expanding population wave, in which the culture spreads with the expansion of people. Ammerman & Cavalli-Sforza (1984) described the wave of advance as “colonization without colonists.”

## CLASSICAL MARKERS AND THE NEOLITHIC TRANSITION

The wave of advance model appeared to be compatible with the rate of spread of the Neolithic measured from radiocarbon dates. More important, however, was the introduction of genetic data into the equation. At the time, these necessarily comprised only the “classical,” non-DNA markers: allele frequencies for blood groups, the tissue antigen HLA system, and some enzymes. Assuming that the Near East and Europe had been relatively isolated during the Upper Palaeolithic and Mesolithic, and had therefore had the opportunity to differentiate genetically, certain predictions were possible that might allow different hypotheses about the spread of agriculture to be distinguished.

A demic diffusion of farmers from the Near East into Europe, involving complete replacement, was thought to predict complete homogeneity of the gene-frequency composition of European and Near Eastern populations (until subsequent differentiation). On the other hand, an entirely cultural spread would leave the two regions with different genetic compositions (assuming that was how they started out). However, a mixed model of demic diffusion involving intermarriage with the Mesolithic population would lead to a gene-frequency cline, or gradient, along the main axis of expansion, with one pole in the Near East and the other in northwest Europe.

It had already been shown by Mourant (1954) that the Rhesus-negative ( $Rh^-$ ) blood-group gene was virtually restricted to Europe, North Africa, and the Near East, with its highest frequency amongst the non-Indo-European-speaking Basques of southwest Europe. Mourant became one of the first biologists to use gene-frequency data to write prehistory when he suggested that this might be because the Basques were a relict of an ancient proto-European population who had mixed with newcomers later on. To Ammerman and Cavalli-Sforza, this idea suggested that the newcomers might have been expanding Near Eastern farmers, spreading agriculture into the continent. They were delighted to discover that the  $Rh^-$  gene was also found at high frequencies in northern and northwest Europe—precisely at the supposed peripheries of the Neolithic expansion.

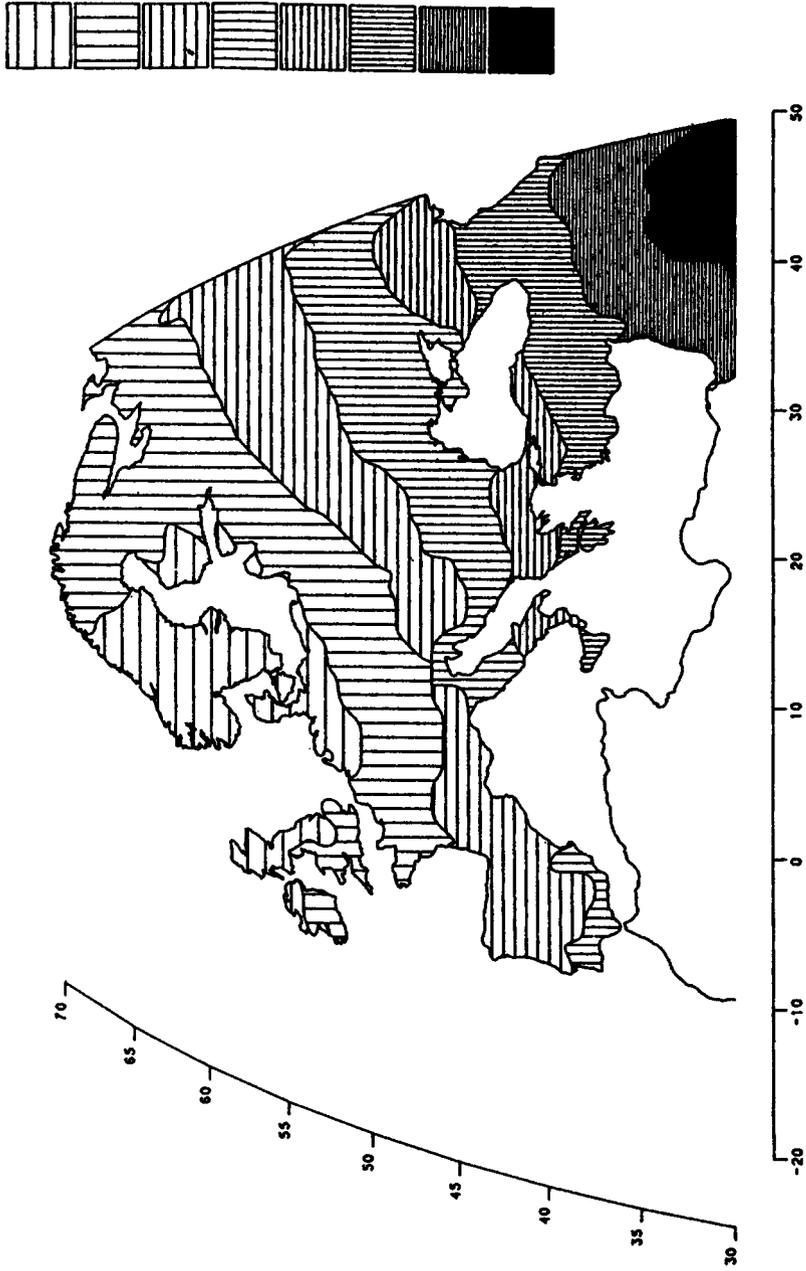
However, other genes often showed different patterns. Furthermore, the Near East and Europe were not in fact highly differentiated from each other, weakening the picture still further. It was therefore necessary to take a multivariate approach, drawing on the results of many genetic systems, and to find an analytical method that could dissect different patterns. Cavalli-Sforza, with colleagues Menozzi and Piazza, chose principal-component (PC) analysis (Menozzi et al. 1978). They used

this to summarize the gene frequencies at a particular location and to represent as much as possible of the information in just a few dimensions. Because the published data sets they used were taken from a variety of sources, there were many gaps, which had to be filled in by interpolation. The results could be graphed as a two-dimensional plot (portraying, say, the first and second or first and third components), or they could be presented as contour maps, component by component, showing the changes in frequency with geography.

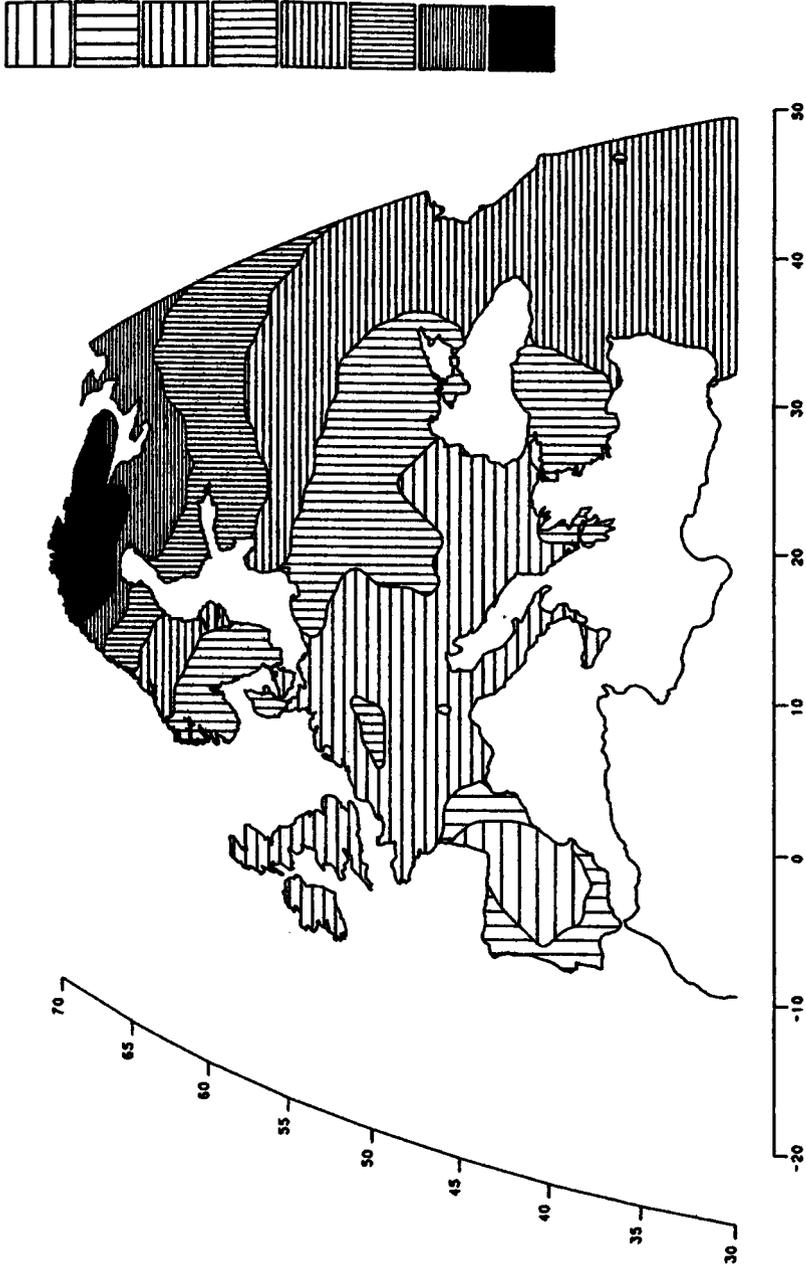
The map of the first PC, accounting for about 27% of the total variation in classical marker frequencies across Europe and the Near East (initially using 39 genetic loci; later 95), has become something of an icon in the archaeogenetics of Europe (Figure 1). It showed a gradient from the southeast to the northwest, with the Near East at one pole and Europe at the other. The resemblance to the radiocarbon map for the spread of the Neolithic was immediately obvious. This was, Cavalli-Sforza and his colleagues believed, strong evidence for the mixed demic diffusion hypothesis. The second and third components (accounting for about 22% and 11% of the variation, respectively) showed clines that were oriented roughly southwest-northeast (Figure 2) and east-west. Because their impact on the genetic variation was lower, the processes that generated these were assumed to have taken place more recently than the Neolithic.

The relative proportions of incoming farmers and indigenous hunter-gatherers were difficult to assess, although the process of diffusion and interaction could be simulated using the wave of advance model (Ammerman & Cavalli-Sforza 1984). The simulations showed that the extent of acculturation—modeled as the marrying of hunter-gatherers into the farming community—was critical. With high levels of intermarriage, the survival of Near Eastern genes in the European population could in fact be very low, even with the wave of advance in operation, and a rapid fall-off of Near Eastern genes toward the northwest would be expected. Thus, whereas genetic data could perhaps be used to assess the extent of demic diffusion, testing the wave of advance model was going to be more difficult.

The conclusions of Ammerman and Cavalli-Sforza and their colleagues were supported by work using a different analytical approach: spatial autocorrelation analysis (Sokal et al. 1989, 1991). Again, about a third of the genetic markers analyzed appeared to be arranged in a southeast-northwest cline. Despite Ammerman and Cavalli-Sforza's early caveats about the acculturation coefficient, and despite the fact that only about a quarter to a third of the variation could be explained by these gradients, the assumed model of surplus-driven population growth and expansion led both groups to tend to play up the role of the Neolithic newcomers at the expense of the indigenous Mesolithic peoples. After all, it was the newcomers who had won in the end. No matter how careful and qualified the argument put forward in 1984 was, by the time of the publication of Cavalli-Sforza's magnum opus, *The History and Geography of Human Genes* (Cavalli-Sforza et al. 1994), the view that the genetic data supported an overwhelmingly Neolithic ancestry for modern Europeans had firmly taken root. It was further reflected in the view that much of the remaining variation was the result of subsequent migrations into



**Figure 1** Synthetic map of the first principal component of variation in 95 classical genetic markers (from Cavalli-Sforza et al., *The History and Geography of Human Genes*, Copyright © 1994 by Princeton University Press).



**Figure 2** Synthetic map of the second principal component of variation in 95 classical genetic markers (from Cavalli-Sforza et al., *The History and Geography of Human Genes*, Copyright © 1994 by Princeton University Press).

Europe and that the principal components could provide relative dating, like an archaeological stratigraphy (Cavalli-Sforza 1996).

Further consequences ensued. Although he was suitably cautious about the genetic evidence for demic diffusion, Renfrew (1987) enthusiastically embraced the wave of advance model. To Renfrew, it was a good processual model that could be rigorously formulated, anti-migrationist (in the traditional sense) and yet allowing for the expansion of a group of people over, potentially, an enormously wide area. It seemed perfectly suited to Renfrew's radical hypothesis that the spread of Indo-European languages through Europe had been mediated by agriculturalist expansion. He suggested that the Basques were indeed a Palaeolithic relict and that the Indo-European languages had been dispersed through Europe by Neolithic populations originating in Anatolia. The choice of Anatolia was necessitated by the fact that this region was home to a number of extinct Bronze-Age Indo-European languages, which were extremely archaic and believed by some linguists to form a sister branch to the surviving members of the family. Many linguists believed them to be intrusive to Anatolia; Renfrew's suggestion was not popular among Indo-Europeanists, who mostly preferred something akin to the hypothesis of Marija Gimbutas that the Proto-Indo-Europeans had spread west with the Bronze-Age "Kurgan culture" of the eastern European steppe (Mallory 1989, Hines 1991). It was, however, received with considerable interest among archaeologists (Zvelebil & Zvelebil 1988, Sherratt & Sherratt 1988).

## CRITIQUES OF THE WAVE OF ADVANCE

Despite some rather vigorous methodological disagreements between the two main proponents of large-scale Neolithic demic diffusion, the Cavalli-Sforza and Sokal groups, the southeast-northwest gradient itself was robust to a number of tests. It appeared there was a genuine pattern that needed explaining, rather than some artefact of the analysis. Nevertheless, interpreting the gradients was not as unproblematic as first thought. This was originally argued not by geneticists, however, but by an archaeologist sympathetic to Renfrew's perspective, Marek Zvelebil.

Firstly, Zvelebil (1989, 1998a) pointed out that there was no strong reason for identifying the first PC solely with a Neolithic expansion. Europe is a small peninsula of the Eurasian landmass and, as such, is likely to have been the sink for many such dispersals throughout prehistory. Movements into Europe may well have taken place many times via Anatolia because this represents one of the main possible points of entry into the continent. The term palimpsest came into play (Renfrew 1998, Zvelebil 1998a): The gradients identified by PC analysis might be the result of many dispersals, each one overwriting the last. The Neolithic may have been one of these, or it may not; if it was, it may or may not have been the most significant. Whereas Zvelebil's argument for "incremental palimpsest" (Zvelebil 2000) focused on post-Neolithic migrations, Richards et al. (1997) compiled a radiocarbon map for the first spread of the Early Upper Palaeolithic into Europe (analogous to the early maps for the Neolithic produced by J.G.D. Clark). They

pointed out that the spread of the first modern humans into Europe, from about 45,000 years ago, followed very similar routes to the later spread of the Neolithic—a rapid dispersal around the Mediterranean and a more gradual expansion along the southeast-northwest axis, following the river systems of central Europe. Whereas any patterns set up by the arrival of early modern humans may well have been erased by subsequent demographic events, this reinforced the point that the same routes into Europe may have been used time and time again.

Thus, the idea of “one PC—one migration,” suggested quite specifically by Cavalli-Sforza, seemed implausible, as did his analogy between principal components and an archaeological stratigraphy. He and his colleagues arranged principal components in temporal succession, according to their magnitude. The first PC, the supposed Neolithic southeast-northwest gradient, came first, accounting for about 27% of the total variation. The third, running east-west, was attributed to Gimbutas’s Proto-Indo-European Kurgan warriors and other invaders from the steppes, and the fourth was ascribed to Greek colonization of the eastern Mediterranean. The second, showing a southwest-northeast gradient and accounting for about 22% of the variation, was thought to be problematic because no plausible demographic explanation in the appropriate time frame (between the early Neolithic and the Bronze Age) could be divined for it. It now seems surprising that the suggestion that the second component might have been in part the result of the Lateglacial re-expansions across Europe from the southwest was not mooted until 1998 (Torroni et al. 1998). Cavalli-Sforza and his colleagues were not ignorant of Palaeolithic processes in Europe, but the fallacious idea that the magnitude of the PCs reflected their age was a blinker to the full range of explanatory possibilities.

Further critiques of the wave of advance have been mounted on archaeological grounds. In the first place, the radiocarbon map of Neolithic spread used by Ammerman & Cavalli-Sforza (1984) was flawed in a number of ways, and the Neolithic package they had used to map the expansion was gradually picked apart. Whereas Ammerman and Cavalli-Sforza took the package as comprising domesticated cereals and animals, ceramics, and so on, various authors showed that these items rarely moved together, except in southeast and central Europe, and that they might often be exchanged into Mesolithic communities (Thomas 1996, Price 2000a). This work suggested that Ammerman and Cavalli-Sforza’s strategy of using the presence of pottery, for example, to identify a settlement as “Neolithic” (adopted to maximize the number of sites included) may have led them to overestimate the Neolithic impact and the uniformity of its spread (Zvelebil 1986). Although Ammerman & Cavalli-Sforza (1984) did emphasize that their rate was merely an average, the whole thrust of the wave of advance model was to impose uniformity on the process. Archaeologists now widely agree that the process was very heterogeneous and that there is no evidence in the archaeological record for large-scale continent-wide immigration (e.g., Pluciennik 1998, Zvelebil 2000). Indeed, detailed studies of northern Europe suggest, in fact, that the Neolithic developed in many areas over a very long period of time and did not arrive in the Baltic region, for example, until well into the Bronze Age or even the Iron Age (Zvelebil 1993, 1998b).

Ammerman and Cavalli-Sforza's coupling of Neolithic populations with high population density and Mesolithic ones with low density has been another casualty of further research. Despite the enormous growth potential of agricultural populations—the rationale for the wave of advance—the archaeological and palynological records suggest that this was never remotely approached during the Neolithic itself. Early Neolithic communities in Europe are likely to have been quite small-scale and mobile (Willis & Bennett 1994; van Andel & Runnels 1995; Roberts 1998, pp. 154–58). By contrast, lacustrine, riverine, and coastal Mesolithic communities are likely to have been home to affluent, complex foraging communities, with much greater population densities and a much higher degree of sedentism than used to be assumed. Ethnographic comparisons with Australian aborigines, or the southern African Khoisan-speaking hunter-gatherers, for example, seem to have been inappropriate for these parts of Mesolithic Europe; comparisons with Northwest Coast Native Americans, for example, now seem more suitable (Zvelebil 1986). In this light, the Mesolithic should not be regarded as a mere prelude to the Neolithic revolution but as a western alternative to it. The situation in the forested interior of central Europe, where the Neolithic is known to have spread extremely rapidly, is much less clear: Mesolithic remains are still scanty in these areas. However, it is curious that in this particular case—in which colonization is almost universally accepted to be the best explanation of the archaeological record (the Bandkeramik or LBK horizon)—the spread was in fact far more rapid than the average rate estimated by Ammerman & Cavalli-Sforza (1984) (Gronenborn 1999, Bogucki 2000, Price 2000b). The same is true for the appearance of Cardial ware in the western Mediterranean, which is also thought likely to represent a case of colonization, in this case by sea (Barnett 2000; Zilhão 2000, 2001).

## MOLECULAR-GENETIC APPROACHES

Since Cavalli-Sforza and his colleagues formulated their views, it has not only been archaeologists who have moved on. There has also, of course, been a revolution in the study of the human genome. In the 1980s, it became possible to analyze not merely the products of certain genes, as had been done in the “classical” analyses, but also the DNA sequences of the genes themselves. In particular, attention has focused on the two non-recombining genetic loci in humans: the mitochondrial DNA (mtDNA), which is present in both sexes but inherited only down the maternal line, and the Y chromosome, which is present only in males and inherited from father to son.

The new molecular data offer two major advantages over the “classical” approach. The first advantage is the estimation of phylogenies or genealogical trees. Both mtDNA and the Y chromosome, because they are non-recombining—so that variation is not reshuffled between lineages—can be seen as genetic systems in which mutations fall onto an independently formed genealogy: the maternal and paternal lines of descent, respectively. Whereas the genealogies of “family history” involve a substantial dose of fiction to draw attention to particular connections

(because the number of ancestors of an individual burgeons alarmingly as one traces them back in time), maternal and paternal ancestry is, by contrast, clearly defined. Any sample of individual subjects will have a defined set of genealogical relations on both the maternal and paternal side so that, in principle, a tree of ancestry could be reconstructed for each. The mtDNA and the Y chromosome both allow us to estimate those trees because both systems have recorded a trace of the pattern of descent, as mutations have inscribed variants into their DNA sequences during the course of history. Because this marks a turn from the collectivist approach of classical population genetics (which necessarily defined “populations” to which allele frequencies were ascribed), it heralds a dramatic increase in resolution for processes involving individuals, such as prehistoric dispersals (Richards & Macaulay 2000). The west Eurasian part of the mtDNA tree is shown in schematic form in Figure 3.

The second major advantage of molecular genetics is dating. Under the neutral theory of molecular evolution, mutations fall onto the genealogy at random and therefore accumulate roughly linearly with time. Just as radiocarbon dating revolutionized archaeology, so the molecular clock has transformed population genetics by providing a timescale. Whether we have reached a stage equivalent to the calibration revolution of radiocarbon in genetics is a subject of some debate. The general consensus is that we haven't; but even so, the introduction of genetic dating does seem to be leading to the collapse of some traditional frameworks for European prehistory in a manner analogous to the famous radiocarbon tree-ring revolution (Renfrew 1973). While always advocating due caution in the interpretation of genetic dates, it is possible to argue that some systems (such as mtDNA) are reasonably well calibrated (Macaulay et al. 1997). The debate does not end here, however. There is also considerable discussion as to how molecular dates can be applied to what is still often seen as the demographic history of populations.

Although many workers have continued to apply (and develop) classical population-genetics methods of analysis to the new molecular data, the new data are different in kind and cry out for new approaches. Of course, one can continue to simply compile frequencies of different forms (or alleles) of a particular gene that are now defined by molecular data. However, the information that this can provide on demographic history is inevitably very impoverished—which is the reason Cavalli-Sforza and his colleagues had to resort to multivariate statistics when analyzing classical markers. By contrast, the non-recombining systems provide us with extremely detailed and fine-grained information about the relations between the alleles, which can be used to reconstruct the genealogical history of the locus. This has led to the development of what has been termed the phylogeographic approach (Richards et al. 1997, Bandelt et al. 2002).

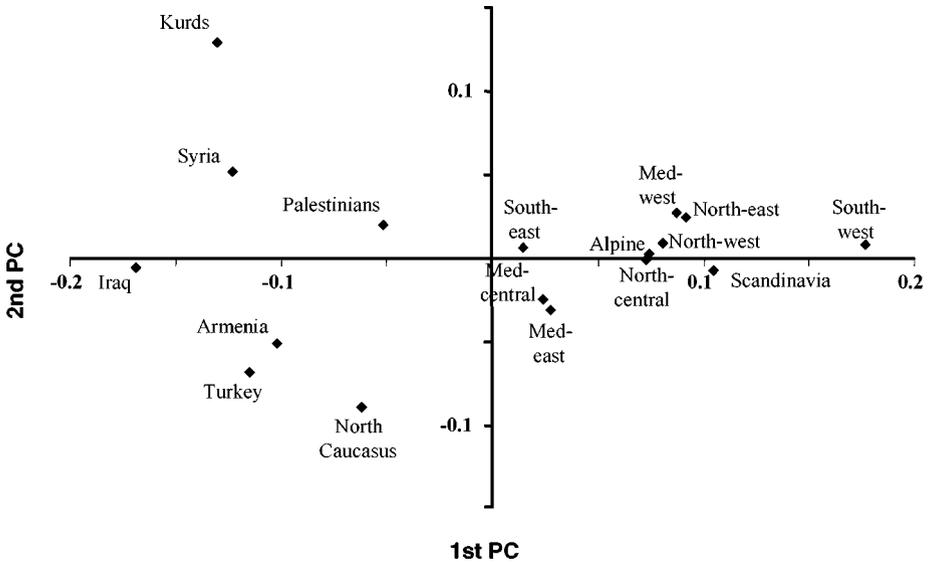
Phylogeography is a heuristic tool for the interpretation of genetic data that attempts to make maximum use of reconstructed trees of descent. It employs the geographic distribution and diversity of genealogical lineages to make inferences about demographic history, in particular, range expansions, migrations, and dispersals. In other words, as Avise (2000, p. 3) has put it, phylogeography is the

mapping of gene genealogies in time and space. To some extent, it can trace its ancestry not only from evolutionary genetics but also from the work of Cavalli-Sforza and his colleagues. They also mapped contemporary human genetic variation in space and attempted to interpret it in terms of prehistoric demography—indeed their approach has often been referred to as “gene geography.” A subdiscipline of biogeography, phylogeography is inherently interdisciplinary, with input from, for example, molecular genetics, population genetics, phylogenetics, demography, climatology, ecology, and historical geography—plus archaeology, anthropology, and linguistics, in the case of humans, and ethology and palaeontology in the case of other species. This interdisciplinarity is obligate: Phylogeographic inferences are (at least at present) generally rather weak and acquire their force when combined in a context built from a number of disciplines. This need not entail circularity (although it is a risk to be wary of). Phylogeographic hypotheses should, in principle, be capable of testing within a preexisting model-based framework (Richards et al. 2002b). This always entails making assumptions, but the assumptions themselves can be susceptible to empirical investigation and may not always be quite as unrealistic as those of more traditional population-genetics approaches (Richards et al. 2000).

## EUROPEAN MITOCHONDRIAL DNA VARIATION

Mitochondrial DNA variation in Europe was studied for the first time by a number of groups in the early 1990s, mostly focusing on sequences from the fast-evolving first hyper-variable segment of the mtDNA control region (HVS-I). Initially, it seemed that the European mtDNA landscape might be so flat as to be almost entirely uninformative with respect to European prehistory (Pult et al. 1994). For example, an attempt by Simoni et al. (2000), using spatial autocorrelation, indicated no southeast-northwest gradient but only a weak gradient along the northern Mediterranean (but see Torroni et al. 2000 for a discussion of flaws hampering this analysis). This finding led to the suggestion that mtDNA may not be a useful demographic marker system, perhaps because of selection or high rates of female gene flow in recent times.

The problem was overcome by supplementing HVS-I data with additional informative variants from the coding region (Torroni et al. 1994, 1996; Macaulay et al. 1999). In this way, mtDNA variation can be dissected into genealogical clades (or haplogroups), which are now strongly supported by analyses of the complete mtDNA genome (Finnilä et al. 2001, Richards & Macaulay 2001) (Figure 3). As shown by Richards et al. (2002a), mtDNA dissected simply into major haplogroups on this basis shows a genetic cline with the Near East at one pole and the Basque country at the other, accounting for 51% of the total variation (Figure 4). The second PC (after subtracting recently migrated African haplogroups) represents the European mtDNAs as a subset of Near Eastern variation. The first PC of mtDNA looks rather similar to the first PC displayed by classical markers: Central and Mediterranean populations group closer to those of the Near



**Figure 4** Plot of the first and second principal components of West Eurasian mtDNA haplogroups (after Richards et al. 2002a).

East, with the remaining European regions between these and the outlier Scandinavians and Basques. How is the pattern to be explained? To address this, much more detailed study of the geographical distribution and time depth of the lineages concerned is necessary. The analysis should not, as in the classical analyses, culminate in a PC plot, from which one reads off the prehistoric processes. With the molecular data, the PC plot is simply an impetus to further study (cf. Clark 1998).

Several novel methodological approaches came together in the analysis of European mtDNA. First, employing a new phylogenetic-network approach to tree reconstruction helped to overcome some of the problems associated with traditional phylogenetic methods, which had plagued earlier work based on human mtDNA (Bandelt et al. 1995). Furthermore, new approaches were being developed to study the colonization of a new area using molecular-genetic data. Finally, the approach pioneered by Cavalli-Sforza himself, of taking the archaeological context seriously, was an essential component.

Founder analysis, an approach for dating the colonization of a region using the molecular clock, had been developed by a number of workers, including Stoneking & Wilson (1989), Torroni et al. (1993), and Forster et al. (1996). The approach works by comparing variation in the region that has been settled (the sink population) with the variation in likely source populations, in order to identify founder types and use them to date individual migration events. This involves simply subtracting from the mutational variation in the sink population the fraction of the variation that arose in the source population and has been carried into the sink region by the founders during the colonization process. Effectively, when a founder

event occurs, owing to an individual migration event from the source to the sink region, the molecular clock is reset: The descendants of that individual will be effectively members of a new lineage tracing to the time of arrival. This is the date that we are interested in because this is the date of the colonization itself. The molecular age of the founder type in the source population will necessarily be older, and perhaps very much older (Richards & Macaulay 2000).

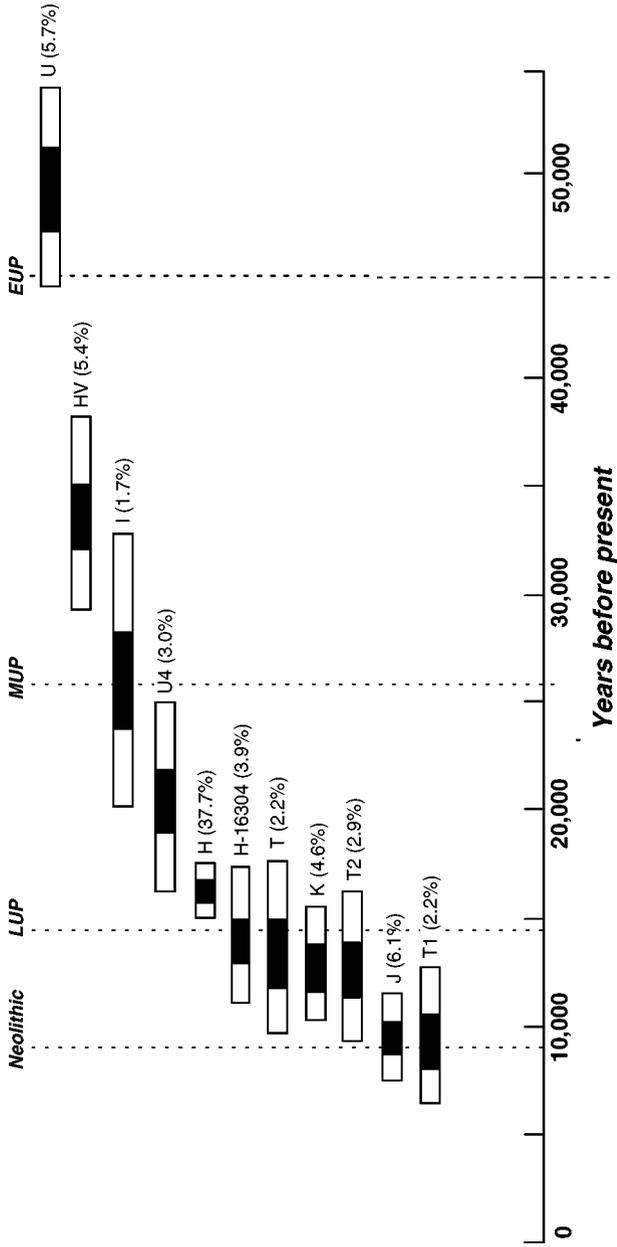
The first results from European mtDNA (Richards et al. 1996, 1998) suggested that only a small minority of lineages dated to the Neolithic, with the remainder dating back to between 15,000 and about 50,000 years ago. The majority appeared to descend from founders of Middle or Late Upper Palaeolithic origin. These clades were strikingly star-like, indicating dramatic population expansions, which suggested that they were mainly the result of re-expansions in the Lateglacial or Postglacial period. The results were, however, rather tentative because they were reliant on comparisons with a very small and inadequate sample from the Near East.

Further work confirmed these outlines, however. Torroni and his colleagues (1998, 2001), focusing on a particular mtDNA clade known as haplogroup V, have shown that there were indeed dramatic Lateglacial expansions from southwest Europe that repopulated much of the continent from about 15,000 years ago. This is supported by recent work on archaeological dates (Housley et al. 1997, Richards et al. 2003). Moreover, as Torroni et al. (1998) pointed out, Lateglacial re-expansions from the southwest provided a plausible explanation for the mysterious second principal component of classical markers, which was oriented southwest-northeast (Figure 2).

In the meantime, Richards et al. (2000) have used a greatly improved Near Eastern mtDNA database, as well as a more sophisticated founder analysis, to quantify the proportions of lineages surviving from the various major dispersal phases in European prehistory. Although it is difficult to extrapolate to the scale of the immigration at the time, it is possible to estimate the proportion of lineages in the modern population that descend from one or another immigration event. Therefore, at least, the question of “farmers our ancestors” could be addressed.

In this analysis, the long-term complexity of interactions between Europe and the Near East, including what appeared to be substantial back-migration from Europe into the Near East, was a major complicating factor (in comparison with analyses of Native Americans, for example). Furthermore, there was the problem of the high rate of mutation in HVS-I, mimicking founder types by recurrent mutation. It was therefore necessary to perform the analysis in a number of different ways that allowed for these complicating factors. In fact, though, the outcomes were reassuringly similar across the different criteria used.

Under the most plausible set of assumptions, three quarters of the modern mtDNA lineages could be traced to just eleven ancestors (the remaining quarter comprising a much larger set of minor founders) (Figure 5). Under various criteria, the putative Neolithic component in modern Europe—that is to say, those lineages that appeared from the Near East about 9000 years ago—occurs at between



**Figure 5** Age ranges for the 11 major founder clades in European mtDNA (those present in 40 or more individuals in a sample of 2804), under the  $f_s$  criterion. The 95% and 50% credible regions for the age estimates are shown by white and black bars. The haplogroup to which the founder clade belongs are indicated, with the proportion of individuals in each in brackets, respectively. For further details, see Richards et al. (2000).

12% and 23%; the best estimate would be around 13%. The Early Upper Palaeolithic component was between 2% and 17%, with a most likely value of about 7% or slightly more. Lateglacial expansions were conflated with preceding Middle Upper Palaeolithic immigration; but, between them, they accounted for about two thirds of modern lineages. It appeared that, on the maternal line of descent, only a minority of European ancestors were Near Eastern farmers; in the main, they were indigenous European hunter-gatherers and foragers, who adopted farming later on.

The detailed analysis of these lineages allows us to interpret the mtDNA PC plot in an informed way. The main contributing clades to the first PC are haplogroups H, pre-V, and U5, concentrated at the European pole, and (pre-HV)1 and U1, concentrated at the Near Eastern pole. The founder analysis suggests that haplogroup U5 appeared in Europe very early, about 45,000 years ago, whereas haplogroups H and pre-V arrived in Europe during the Middle Upper Palaeolithic and re-expanded after the Last Glacial Maximum. Their geographical distributions support this scenario. Haplogroups U5 and V are predominantly European, whereas haplogroups (pre-HV)1 and U1 are predominantly Near Eastern, with some probably recent gene flow along the Mediterranean. Haplogroup H is the most frequent cluster in both Europe and the Near East, but it occurs at frequencies of only around 25%–30% in the Near East; whereas its frequency is about 45%–50% in most European populations, reaching about 60% in the Basque country. It is most frequent in western Europe and least frequent in the Near East: quite the opposite of what one would expect had it been distributed by demic diffusion during the Neolithic. Indeed, the founder analysis suggests that the main Neolithic founders are likely to have been members of haplogroups J and T1; neither of these is a determining factor in the first PC.

Furthermore, the pattern of the first PC of mtDNA indicates similarity between Europeans and Near Easterners primarily along the Mediterranean, whereas archaeological evidence would suggest that the main expansion is likely to have been into the Balkans and thence into central Europe. The founder analysis of mtDNA suggests that the similarities along the Mediterranean may be due, at least in part, to substantial recent gene flow rather than solely to the Neolithic expansion. The (pre-HV)1 clade, however, may well have spread along the Mediterranean during the Neolithic. This idea suggests that the first PC of mtDNA is indeed a palimpsest, witness to some processes more recent than the Neolithic and some more ancient.

## CRITICISMS OF THE MITOCHONDRIAL DNA WORK

The mtDNA work has been criticized from a traditional population-genetics perspective by a number of authors (Cavalli-Sforza & Minch 1997; Barbujani et al. 1998; Chikhi et al. 1998, 2002; Barbujani & Bertorelle 2001). The common thread in these critiques is that “the age of a population is not the age of the common molecular ancestor of its set of DNA sequences” (Barbujani & Chikhi 2000), with

the implication being that this point has been somehow missed in the analysis. As we have seen, however, this is hardly the case: Founder analysis was explicitly designed to get around this problem (Richards et al. 2000, Richards & Macaulay 2000).

Chikhi et al. (1998) developed a dating approach based on population splits. They used spatial autocorrelation to show that geographical gradients were present not only in classical markers but also in seven autosomal molecular marker systems. Although they did not identify the direction of these clines, they interpreted them as indicating a directional expansion. Of course, simply identifying clines does not tell us anything about when they were generated. More specifically, however, on the evidence of population divergence times estimated from four microsatellite loci, they argued that the clines were the result of Neolithic demic diffusion from the Near East into Europe. This argument was made on the grounds that most of the population splits were estimated to be very recent. In fact, the great majority appeared to be less than 1000 years—not Neolithic at all—and splits involving the supposedly pre-Neolithic Basques did not feature among the very few ancient splits identified, which mostly involved the Saami. However, Saami mtDNA lineages have undergone very severe founder effects, and it is this feature, rather than some innate Palaeolithic qualities that other Europeans lack, that causes them to stand out (Torroni et al. 1998). Such an approach should, at least, allow for regions where the Neolithic is most likely to have been indigenous, such as Iberia and Scandinavia, to act as controls; at the least, these regions should indicate high time depths. But this was not the case in these analyses, which ignore the archaeological evidence in an attempt to read the demographic history directly from the genetic-distance matrices.

In fact, similar results would be obtained from the mtDNA data. Genetic distances can only be used to estimate the time to population fissions under very particular circumstances: The ancestral and daughter populations have to be constant and equal in size, at least for relatively recent events, with no subsequent gene flow. These conditions are hardly likely to be remotely approached for human groups, which have usually undergone founder effects and expansions through time, not to mention multiple fissions and fusions and frequent exchange of genes. As Bandelt et al. (2002) have shown, a similar approach would identify both the Basque and Korean gene pools as having been formed as a result of Neolithic demic diffusion from the Near East.

A subsequent analysis (Barbujani & Bertorelle 2001), this time using haplotype matching to estimate population divergences, similarly placed most of them in the Bronze Age (hence interpreted as supporting Neolithic demic diffusion). It is curious that this analysis relied on a model of exponential expansion over the last 40,000 years (again ignoring archaeological evidence to the contrary), rather than constant size as in the previous analysis, in order to draw the same conclusion (Bandelt et al. 2002).

Perhaps the most important question for the founder analysis concerns the statistical validity of the results, since the analysis relies on the sample size in

the source population being adequate to identify all of the most important founder types. Although the founder analysis described here was not sufficiently automated to allow for a resampling test (such as the bootstrap), Richards et al. (2000) did perform a reanalysis in which half of the source data were omitted. This gave very similar results (particularly for the Neolithic contribution) to the analyses using the full source data. This reanalysis uses only the core Fertile Crescent data, omitting Anatolia, Egypt, and the southern Caucasus. It may also help, therefore, to address the Eurocentric bias of the main analyses, drawing as they do sharp dividing lines between “Europe” and the “Near East” at the Bosphorus and Caucasus mountains (M. Özdoğan, personal communication).

## REGIONAL VARIATION IN EUROPE

The founder analysis as described above dates the arrival of mtDNAs in Europe as a whole [or, in the case of the Lateglacial expansions, the timing of a bottleneck within Europe itself (Richards et al. 2000, 2003)]. As mentioned earlier, over the past 25 years archaeology has become much more interested in the regional variation rather than the overall process. Richards et al. (2000) therefore repeated the analysis at the regional level. It has to be stressed that this approach has several limitations. First, the results for any one region are based on fewer data and are therefore naturally associated with greater uncertainty. Second, the data are of variable quality and may poorly represent the deep ancestry of lineages within each region in some cases, such as eastern Europe and Greece. Finally, the results are, at best, estimating the proportion of lineages in the present-day population that can be attributed to each founder event from the Near East (or to bottlenecks within Europe), rather than from the immediate source region. The results are worth discussing, however, as there are both interesting congruences and some conflicts with the archaeological patterns.

The highest Neolithic impact, according to this analysis, was on southeast Europe, central Europe, and northwest and northeast Europe, which showed values of between 15% and 22% Neolithic lineages each. The Neolithic lineages are mainly from haplogroup J and include a specific subset of J lineages, called J1a, that are restricted to this region and seem to be a marker for the LBK (*Linienbandkeramische Kultur*) and post-LBK dispersals (Richards et al. 1996). For southeast and central Europe, a relatively high Neolithic component is very much in agreement with the usual interpretation of the archaeological record. There is some consensus that both the Balkans Neolithic and the central European LBK were the result of direct colonization—although there is debate about the extent to which the process may have involved acculturation as well (cf. Gronenborn 1999, Tringham 2000, Budja 2001). Acculturation may also have taken place in between the two processes, where there was a substantial break in the expansion (Bogucki 2000, Zvelebil 2000). The mtDNA results suggest that colonization from (ultimately) the Near East did indeed take place and that the descendants of Near Eastern

colonists are represented in the central European populations of the present day. Nevertheless, more than three quarters of the surviving lineages are the result of acculturation of indigenous foraging peoples. This finding appears to support the integrationist model described by Zvelebil (2000), which involves both pioneer “leapfrog” colonization (directed toward suitable land) and acculturation and genetic exchange across the agricultural frontier during the phase in which aspects of farming become available to the surrounding foraging populations. Strontium isotope analysis has recently suggested immigration, into LBK settlements from very early times (Bentley et al. 2002), of nonlocal people who may have been brought in from the foraging communities (Gronenborn 1999).

The presence of Near Eastern lineages at similar frequencies in the northwest, however, seems to conflict with both Zvelebil’s model and the patterns of the classical markers and the Y chromosome (see below). Zvelebil suggests that a long-term frontier was established on the north European plain and that the transition to farming to the north, northwest, northeast, and southwest took place largely by acculturation. The mtDNA picture, however, suggests a value of about 20% for Near Eastern lineages in the northwest (the northeast is more equivocal because the sample may be insufficiently representative). By contrast, both classical markers and the Y chromosome indicate few or no Neolithic markers in the British Isles. In this case, it is perhaps possible that there were female-only exchanges between the post-LBK peoples of the North European plain and the northwest across the agricultural frontier (Wilson et al. 2001). Alternatively, there may have been acculturation at the LBK frontier, after which (by chance) predominantly Near Eastern mtDNAs but predominantly acculturated Y chromosomes moved on to the northwest (Renfrew 2001). Finally, it is possible that the mtDNA lineages were dispersed into the northwest by later processes.

There are fewer Neolithic-derived lineages along the Mediterranean and the Atlantic west, at around 10%, again mainly from haplogroup J. The sample from the eastern Mediterranean is rather small and poorly provenanced but would be compatible with maritime colonization of Greece by Near Eastern pioneer groups (Perlès 2001). Again, there appear to be some mtDNAs that are regionally specific, such as J1b1 along the Atlantic facade. This brings to mind the archaeological view of some colonization alongside acculturation of quite dense, sedentary Mesolithic communities (Barnett 2000; Zilhão 2000, 2001). The Basque region, which was an outlier in the PC analyses of both mtDNA and classical markers, has the lowest Neolithic component, at around 7%. The Basque outlier status may therefore be partly the result of reduced Neolithic penetration, as well as considerable genetic drift due to isolation and small population size. They are little more of a Mesolithic relict than any other European population.

A striking, if perhaps unsurprising, result of the analysis was that there have been very high levels of more recent gene flow in the eastern Mediterranean. In fact, in Greece approximately 20% of lineages have a recent Near Eastern origin. Similarly, in many parts of the Near East, back-migration of lineages from Europe is estimated to be approximately 20%.

## Y-CHROMOSOME VARIATION AND THE NEOLITHIC

Around the same time as the early mtDNA work was being published, a number of studies were also attempting to dissect Paleolithic and Neolithic Y-chromosome lineages within Europe. Semino et al. (1996) identified candidates both for an indigenous European clade of lineages, now known as paragroup R\* (in the nomenclature of the Y Chromosome Consortium 2002; see also Hammer & Zegura 2002), and a likely Near Eastern Neolithic component, now called haplogroup J (by analogy with the most common putative Neolithic marker in the mtDNA). The Y-chromosome haplogroup J showed a cline similar to Cavalli-Sforza's first PC for the classical markers, whereas R\* was most common in western Europe and declined moving east. The highest diversity of haplogroup J appears to be in the Fertile Crescent, possibly Iran (Quintana-Murci et al. 2001).

The work has been developed in more genealogical detail by Semino et al. (2000) (Figure 6) using a large series of new markers discovered by Underhill et al. (2000), with fewer markers but many more samples by Rosser et al. (2000), and with greater focus by Scozzari et al. (2001). These studies confirm that the most common diagnostic Near Eastern haplogroups are haplogroup J, referred to above, and haplogroup E3b. Both J and E3b display declining gradients moving from the Near East to Europe, and both may have potentially been spread with the Neolithic. This would imply a Near Eastern Neolithic contribution to Europe, as a whole, of about 20%–25%, similar to the estimates from mtDNA (Semino et al. 2000).

This work has been recently criticized by Chikhi et al. (2002), who have re-analyzed the Y-chromosome data set of Semino et al. (2000) using an admixture approach. They suggested that the Neolithic contribution is much higher—in fact, greater than 50% in most parts of Europe, rather than <25% overall. However, to claim this they have had to assume that the various parts of Europe were formed by a unitary Palaeolithic component (represented by modern Basques or Sardinians), admixing with a unitary Near Eastern component (represented by the modern Near East). No allowance was made for back-migration into the Near East, which the mtDNA data shows has been considerable (Richards et al. 2000). Even when allowing for the implausibility of these assumptions, their approach does not avoid the pitfall of the original classical analyses: the lack of a time scale. Therefore, it seems very unlikely that their analysis improves on the more straightforward approach taken by Semino and her colleagues.

Strangely, however, both of the putative Neolithic Y-chromosome haplogroups are more common along the Mediterranean than in central Europe. This picture differs somewhat both from the mtDNA picture described above and the usual understanding of the archaeological evidence, in which the most likely parts of Europe to have been colonized by substantial numbers of Neolithic immigrants are often thought to be southeast and central Europe. Moreover, without a founder analysis, it is possible that earlier and later processes may be conflated. For example, the mtDNA results suggest substantial very recent gene flow from Turkey to

Greece; this would appear as Neolithic in the calculations of Semino et al. (2000). This is the palimpsest problem all over again.

The issue has been addressed in part by King & Underhill (2002), not by means of a founder analysis (which would require more Y-chromosome data than presently exists) but by using an innovative combination of archaeology and genetics. They have compiled a database of painted pottery and anthropomorphic clay figurines from the Near East. These are thought to have originated in the Pre-Pottery Neolithic B (PPNB) phase of the Levantine and Anatolian Neolithic and to have been carried westward with the Neolithic into parts of Europe, including Greece and the Balkans, as far as the Danube basin and the central Mediterranean coastline. King & Underhill tested the geographical distribution of these artefacts for correlation with the distribution of Y-chromosome lineages. Haplogroup E3b (referred to by them as Eu4) and one subclade within haplogroup J (their Eu9) correlated strongly with the distribution of the archaeological markers. Another subclade within haplogroup J correlated with figurines only, which are more widely dispersed than the painted pottery. However, the western Mediterranean displayed the Near Eastern Y-chromosome lineages but no archaeological correlate. King & Underhill therefore suggested that these latter areas might indeed represent more recent gene flow from the eastern Mediterranean, such as Greek or Phoenician colonization. In cases where the correlation held, they proposed that the Near Eastern lineages could indeed be taken as a signal of Neolithic dispersals or of some other kind of demographic interactions with the PPNB Near East. Furthermore, at least some of the Near Eastern Y-chromosome types in the areas of Neolithic artefacts are also likely to have arrived more recently, and perhaps also even earlier, than the Neolithic (cf. also Malaspina et al. 2001). However, it is also possible that Neolithic colonization of the Mediterranean from the Near East involved maritime pioneers who were predominantly male and that may also help to explain the much higher male contribution of Neolithic lineages in the east and central Mediterranean (Perlès 2001). Overall, it seems fair to say that the mtDNA and the Y chromosome do indeed appear to be moving toward a consensus, consistent with the archaeological evidence, on the scale of Neolithic ancestry in modern Europe (Lell & Wallace 2000).

## MODES OF DISPERSAL

It appears, therefore, that the overall Neolithic contribution to modern Europeans is somewhere between 12% and 23% on the female side, with the most likely value being about 13%. It is probably somewhat less than 22% on the male side, depending on how much overwriting there has been in recent times with Near Eastern lineages in southern Europe. From both perspectives, it is clear that the ancestry of the majority of lineages predates the Neolithic in Europe, stretching back to the Last Glacial Maximum and beyond. What is true of the non-recombining marker systems is likely to be true for the autosomal genes as well, and indeed it seems that

this may not be inconsistent with the PC maps of classical markers (Cavalli-Sforza & Minch 1997).

What about the demographic mechanisms involved with the arrival of the Neolithic? Renfrew (2001) has pointed out that even the rapid decline in Near Eastern lineages witnessed on the Y chromosome as one crosses into Europe is not inconsistent with the wave of advance model—although the mtDNA pattern is difficult to reconcile with it. But the range of possible models to consider is now much greater than in 1984. Zvelebil (2000) has listed seven possible mechanisms. These are:

1. folk migration, the traditional migrationist explanation: the directional movement of a whole population from one region to another, leading to genetic replacement;
2. demic diffusion by means of a wave of advance;
3. elite dominance, in which a social elite penetrates an area and imposes a new culture on the local population;
4. infiltration of a community by small numbers of specialists fulfilling a particular need, such as livestock farmers;
5. leapfrog colonization by small groups targeting optimal areas to form an enclave surrounded by indigenous inhabitants;
6. frontier mobility, or exchange between farmers and foragers at agricultural frontier zones; and
7. regional contact, involving trade and exchange of ideas.

Clearly, item 1 would involve genetic replacement—the classic migrationist position. Item 7 would involve no movement of genes whatsoever—referred to by some as indigenism (Ammerman 1989). Indigenism includes both cultural diffusion, championed in the 1980s by Dennell (1983), Barker (1985), and Whittle (1996), and a position more akin to separate development, in which the social and ideological, rather than economic, aspects of the Neolithic are regarded as central (e.g., Hodder 1990; Thomas 1996, 1998).

However, items 2–6 would all involve the arrival of new genetic lineages in an area and the eventual acculturation of the indigenous communities: Zvelebil's integrationism. Elite dominance would likely show little evidence of the newcomers and might be unlikely in any case to be relevant in the early Neolithic (Renfrew 1987). The wave of advance model would predict continent-wide genetic clines, as Ammerman & Cavalli-Sforza (1984) showed. There are clines in some classical markers, but those on the Y chromosome believed to associate specifically with the Neolithic fall off steeply, and clines are not evident in the putative Neolithic mtDNAs. Infiltration and leapfrog colonization would be likely to leave traces in the regions where they had occurred but not in the form of clear clines. Frontier mobility would allow for genetic exchange between colonized, newly Neolithic areas such as central Europe and forager strongholds to the north and west. This pattern indeed fits the mtDNA evidence but not exactly in the way that Zvelebil

predicted. The northwest of Europe has virtually equilibrated with central Europe with respect to its maternal Neolithic lineages, whereas Iberia and Scandinavia do indeed have a reduced proportion of Neolithic mtDNAs.

Overall, the genetic evidence seems most consistent with pioneer leapfrog colonization of southeast and central Europe, with subsequent infilling acculturation of much larger numbers of indigenous foragers. If there was a wave of advance during the rapid expansion in the LBK area, it is more likely to have involved largely mtDNA and Y-chromosome lineages from assimilated Balkan foraging populations, rather than from the Near East. There seems to have been further colonization from the LBK zone into the northwest, including mainland Britain, but the pattern in Scandinavia might be explained by frontier exchange. Much of the Atlantic west (including western Britain) may also have received mtDNAs by frontier exchange, but the west seems also to have experienced distinct leapfrog colonization events from the direction of the Mediterranean coastline (presumably by boat). The movements into the northwest seem either to not have involved men or (perhaps more plausibly) to have involved male lineages that had undergone acculturation and were therefore indigenous to central Europe. In all or most regions of Europe, there seems to have been local adoption of agriculture. This implies that if, as Renfrew (1987) suggested, the Indo-European languages were also dispersed at this time, then contact-induced language change must have been an important mechanism (Zvelebil 1995b).

It is, of course, rather speculative to try and read a detailed demographic picture from the distribution of present-day genetic lineages. It should perhaps be stressed that nothing intrinsically associates any particular mtDNA, or Y chromosome, with the spread of the Neolithic. These reconstructions are made on the basis of the estimated time of arrival of particular lineages and their geographical distribution; and alternative accounts are certainly possible (Bandelt et al. 2002). For example, it is possible that the initial wave of Neolithic expansion in Europe was entirely the result of cultural diffusion and that it was followed by the lineages we now regard as Neolithic as new social networks were established and men and women moved along them (cf. Budja 2001). Any reconstructions suggested today are likely to become as outmoded as the wave of advance as more detailed accounts become possible by combining regional archaeological and genetic information at an ever-increasing level of resolution. At the more general level, though, it is possible to claim there has been a major shift in opinion on the ancestry of Europeans, driven largely by the combined forces of archaeology and genetics. Near Eastern farmers played their part, but the majority of European genetic lineages have their roots in the European Palaeolithic.

## ACKNOWLEDGMENTS

I am very grateful to Hans-Jürgen Bandelt, Detlef Gronenborn, Vincent Macaulay, and Alison Richards for reading and making valuable comments on the text. I thank Princeton University Press for permission to reproduce Figures 1 and 2 and Ornella Semino for permission to use and adapt Figure 6.

The *Annual Review of Anthropology* is online at <http://anthro.annualreviews.org>

## LITERATURE CITED

- Ammerman AJ. 1989. On the Neolithic transition in Europe: a comment on Zvelebil and Zvelebil (1988). *Antiquity* 63:162–65
- Ammerman AJ, Cavalli-Sforza LL. 1984. *The Neolithic Transition and the Genetics of Populations in Europe*. Princeton, NJ: Princeton Univ. Press
- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–65
- Avise JC. 2000. *Phylogeography*. Cambridge, MA: Harvard Univ. Press
- Bandelt H-J, Forster P, Sykes BC, Richards MB. 1995. Mitochondrial portraits of human populations using median networks. *Genetics* 141:743–53
- Bandelt H-J, Macaulay VA, Richards MB. 2002. What molecules can't tell us about the spread of languages and the Neolithic. In *Origins and Dispersals of Agricultural Societies and Language Families*, ed. C Renfrew, P Bellwood, pp. 99–107. Cambridge, UK: McDonald Inst. Archaeol. Res.
- Barbujani G, Bertorelle G. 2001. Genetics and the population history of Europe. *Proc. Natl. Acad. Sci. USA* 98:22–25
- Barbujani G, Bertorelle G, Chikhi L. 1998. Evidence for Paleolithic and Neolithic gene flow in Europe. *Am. J. Hum. Genet.* 62:488–91
- Barbujani G, Chikhi L. 2000. Genetic population structure of Europeans inferred from nuclear and mitochondrial DNA polymorphisms. See Renfrew & Boyle 2000, pp. 119–29
- Barker G. 1985. *Prehistoric Farming in Europe*. Cambridge, UK: Cambridge Univ. Press
- Barnett WK. 2000. Cardial pottery and the agricultural transition in Mediterranean Europe. See Price 2000c, pp. 93–116
- Bentley RA, Price TD, Luning J, Gronenborn D, Wahl J, Fullagar PD. 2002. Prehistoric migration in Europe: strontium isotope analysis of early Neolithic skeletons. *Curr. Anthropol.* 43:799–804
- Binford LR. 1968. Post-Pleistocene adaptations. In *New Perspectives in Archaeology*, ed. SR Binford, LR Binford, pp. 313–41. Chicago: Aldine
- Bogucki P. 2000. How agriculture came to north-central Europe. See Price 2000c, pp. 197–218
- Bradley R. 1984. *The Social Foundations of Prehistoric Britain*. New York: Longman
- Budja M. 2001. The transition to farming in Southeast Europe: perspectives from pottery. In *Documenta Praehistorica XXVIII: 8th Neolithic Studies*, ed. M Budja, pp. 27–47. Ljubljana, Slovenia: Zalozila
- Cavalli-Sforza LL. 1996. The spread of agriculture and nomadic pastoralism. See Harris 1996, pp. 51–69
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. *The History and Geography of Human Genes*. Princeton, NJ: Princeton Univ. Press
- Cavalli-Sforza LL, Minch E. 1997. Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 61:247–51
- Chikhi L, Destro-Bisol G, Bertorelle G, Pascali V, Barbujani G. 1998. Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proc. Natl. Acad. Sci. USA* 95:9053–58
- Chikhi L, Nichols RA, Barbujani G, Beaumont MA. 2002. Y genetic data support the Neolithic demic diffusion model. *Proc. Natl. Acad. Sci. USA* 99:11,008–13
- Childe VG. 1925. *The Dawn of European Civilization*. London: Kegan Paul
- Clark GA. 1998. Multivariate pattern searches, the logic of inference, and European prehistory: a comment on Cavalli-Sforza. *J. Anthropol. Res.* 54:406–11
- Clark JGD. 1965. Radiocarbon dating and the expansion of farming culture from the Near

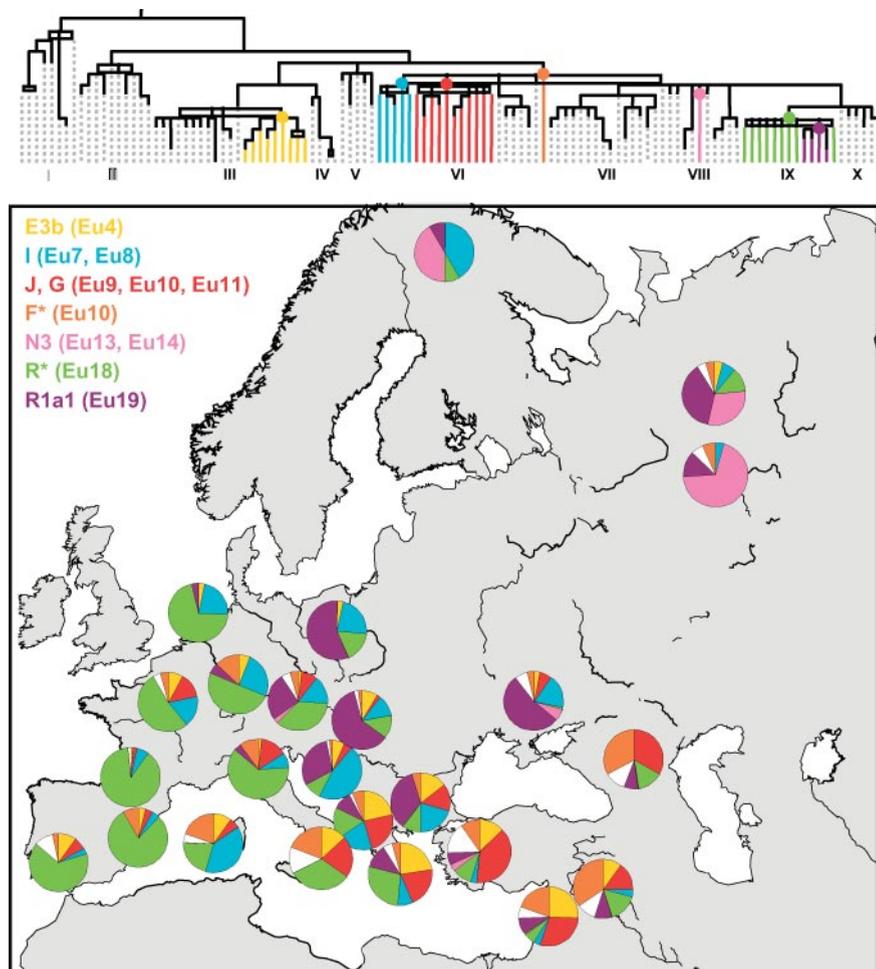
- East over Europe. *Proc. Prehist. Soc.* 31:57–73
- Dennell R. 1983. *European Economic Prehistory: a New Approach*. London: Acad. Press
- Edmonds M, Richards C, eds. 1998. *Understanding the Neolithic of North-Western Europe*. Glasgow, UK: Cruithne Press
- Finnilä S, Lehtonen MS, Majamaa K. 2001. Phylogenetic network for European mtDNA. *Am. J. Hum. Genet.* 68:1475–84
- Forster P, Harding R, Torroni A, Bandelt H-J. 1996. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.* 59:935–45
- Gamble C. 1992. Uttermost ends of the earth. *Antiquity* 66:252–310
- Gronenborn D. 1999. A variation on a basic theme: the transition to farming in southern Central Europe. *J. World Prehist.* 2:23–210
- Hammer MF, Zegura SL. 2002. The human Y chromosome haplogroup tree: nomenclature and phylogeography of its major divisions. *Annu. Rev. Anthropol.* 31:303–21
- Harris DR, ed. 1996. *The Origins and Spread of Agriculture and Pastoralism in Eurasia*. London: UCL Press
- Hines J. 1991. Reviews. *Nor. Archaeol. Rev.* 24:49–54
- Hodder I. 1990. *The Domestication of Europe*. Oxford, UK: Blackwell
- Housley RA, Gamble CS, Street M, Pettitt P. 1997. Radiocarbon evidence for the Lateglacial human recolonisation of northern Europe. *Proc. Prehist. Soc.* 63:25–54
- King R, Underhill P. 2002. Congruent distribution of Neolithic painted pottery and ceramic figurines with Y-chromosome lineages. *Antiquity* 76:707–14
- Kohl PL, Fawcett C. 1995. *Nationalism, Politics and the Practice of Archaeology*. Cambridge, UK: Cambridge Univ. Press
- Lell JT, Wallace DC. 2000. The peopling of Europe from the maternal and paternal perspectives. *Am. J. Hum. Genet.* 67:1376–81
- Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, et al. 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* 64:232–49
- Macaulay VA, Richards MB, Forster P, Bendall KA, Watson E, et al. 1997. mtDNA mutation rates—no need to panic. *Am. J. Hum. Genet.* 61:983–86
- Malaspina P, Tsopanomichalou M, Duman T, Stefan M, Silvestri A, et al. 2001. A multistep process for the dispersal of a Y chromosomal lineage in the Mediterranean area. *Ann. Hum. Genet.* 65:339–49
- Mallory JP. 1989. *In Search of the Indo-Europeans*. London: Thames and Hudson
- Menozzi P, Piazza A, Cavalli-Sforza LL. 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201:786–92
- Mourant AE. 1954. *The Distribution of the Human Blood Groups*. Oxford, UK: Blackwell
- Perlès C. 2001. *The Early Neolithic in Greece*. Cambridge, UK: Cambridge Univ. Press
- Pluciennik M. 1998. Deconstructing the Mesolithic–Neolithic transition. See Edmonds & Richards 1998, pp. 61–83
- Price TD. 2000a. Europe's first farmers: an introduction. See Price 2000c, pp. 1–18
- Price TD. 2000b. Lessons in the transition to agriculture. See Price 2000c, pp. 301–18
- Price TD, ed. 2000c. *Europe's First Farmers*. Cambridge, UK: Cambridge Univ. Press
- Pult I, Sajantila A, Simanainen J, Georgiev O, Schaffner W, Pääbo S. 1994. Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biol. Chem. Hoppe-Seyler* 375: 837–40
- Quintana-Murci L, Krausz C, Zerjal T, Sayer SH, Hammer MF, et al. 2001. Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *Am. J. Hum. Genet.* 68:537–42
- Renfrew C. 1973. *Before Civilization*. London: Penguin
- Renfrew C. 1987. *Archaeology and Language*. Harmondsworth, UK: Penguin
- Renfrew C. 1998. Comments on Cavalli-Sforza and Otte. *J. Anthropol. Res.* 54:417–19
- Renfrew C. 2001. From molecular genetics to

- archaeogenetics. *Proc. Natl. Acad. Sci. USA* 98:4830–32
- Renfrew C, Boyle K, eds. 2000. *Archaeogenetics: DNA and the Population Prehistory of Europe*. Cambridge, UK: McDonald Inst.
- Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, et al. 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 59:185–203
- Richards M, Gamble C, Davies W, Pettitt P. 2003. Upper Palaeolithic refugia: genetic and archaeological perspectives. In *The Genetic Revolution and Archaeology*, ed. C Gosden, M Richards, R Ward. Oxford, UK: Oxford Univ. Press. In press
- Richards M, Macaulay V. 2000. Genetic data and the colonization of Europe: genealogies and founders. See Renfrew & Boyle 2000, pp. 139–51
- Richards M, Macaulay V. 2001. The mitochondrial gene tree comes of age. *Am. J. Hum. Genet.* 68:1315–20
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, et al. 2000. Tracing European founder lineages in the Near Eastern mitochondrial gene pool. *Am. J. Hum. Genet.* 67:1251–76
- Richards M, Macaulay V, Sykes B, Pettitt P, Hedges R, et al. 1997. Reply to Cavallisforza and Minch. *Am. J. Hum. Genet.* 61:251–54
- Richards M, Macaulay V, Torroni A, Bandelt H-J. 2002a. In search of geographical patterns in European mtDNA. *Am. J. Hum. Genet.* 71:1168–74
- Richards MB, Macaulay VA, Bandelt H-J. 2002b. Analyzing genetic data in a model-based framework: inferences about European prehistory. In *Origins and Dispersals of Agricultural Societies and Language Families*, ed. C Renfrew, P Bellwood. Cambridge, UK: McDonald Inst. Archaeol. Res. pp. 459–66
- Richards MB, Macaulay VA, Bandelt H-J, Sykes BC. 1998. Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* 62:241–60
- Roberts N. 1998. *The Holocene*. Oxford, UK: Blackwell
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, et al. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than language. *Am. J. Hum. Genet.* 67:1526–43
- Scozzari R, Cruciani F, Pangrazio A, Santolamazza P, Vona G, et al. 2001. Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. *Hum. Immunol.* 62:871–84
- Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti S. 1996. A view of the neolithic demic diffusion in Europe through two Y chromosome-specific markers. *Am. J. Hum. Genet.* 59:964–68
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, et al. 2000. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290:1155–59
- Sherratt A, Sherratt S. 1988. The archaeology of Indo-European: an alternative view. *Antiquity* 62:584–95
- Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani B. 2000. Geographic patterns of mtDNA diversity in Europe. *Am. J. Hum. Genet.* 66:262–78
- Sokal RR, Harding RM, Oden NL. 1989. Spatial patterns of human gene frequencies in Europe. *Am. J. Phys. Anthropol.* 80:267–94
- Sokal RR, Oden NL, Wilson C. 1991. Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* 351:143–45
- Stoneking M, Wilson AC. 1989. Mitochondrial DNA. In *The Colonization of the Pacific: a Genetic Trail*, ed. AVS Hill, S Serjeantson, pp. 215–45. Oxford, UK: Oxford Univ. Press
- Thomas J. 1996. The cultural context of the first use of domesticates in continental Central and Northwest Europe. See Harris 1996, pp. 310–22
- Thomas J. 1998. Towards a regional geography of the Neolithic. See Edmonds & Richards 1998, pp. 37–60
- Torroni A, Bandelt H-J, D'Urbano L, Lahermo

- P, Moral P, et al. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* 62:1137–52
- Torroni A, Bandelt H-J, Macaulay V, Richards M, Cruciani F, et al. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am. J. Hum. Genet.* 69:844–52
- Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, et al. 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144:1835–50
- Torroni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC. 1994. mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am. J. Hum. Genet.* 55:760–76
- Torroni A, Richards M, Macaulay V, Forster P, Villems R, et al. 2000. mtDNA haplogroups and frequency patterns in Europe. *Am. J. Hum. Genet.* 66:1173–77
- Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, et al. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.* 53:563–90
- Trigger BG. 1989. *A History of Archaeological Thought*. Cambridge, UK: Cambridge Univ. Press
- Tringham R. 2000. Southeastern Europe in the transition to agriculture to Europe: bridge, buffer, or mosaic. See Price 2000c, pp. 19–56
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, et al. 2000. Y chromosome sequence variation and the history of human populations. *Nat. Genet.* 26:358–61
- van Andel TH, Runnels CN. 1995. The earliest farmers in Europe. *Antiquity* 68:481–500
- Whittle A. 1996. *Europe in the Neolithic*. Cambridge, UK: Cambridge Univ. Press
- Willis KJ, Bennett KD. 1994. The Neolithic transition—fact or fiction? Palaeoecological evidence from the Balkans. *The Holocene* 4:326–30
- Wilson JF, Weiss DA, Richards M, Thomas MG, Bradman N, Goldstein DB. 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc. Natl. Acad. Sci. USA* 98:5078–83
- The Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* 12:339–48
- Zilhão J. 2000. From the Mesolithic to the Neolithic in the Iberian peninsula. See Price 2000c, pp. 144–82
- Zilhão J. 2001. Radiocarbon evidence for maritime pioneer colonisation at the origins of farming in west Mediterranean Europe. *Proc. Natl. Acad. Sci. USA* 98:14,180–85
- Zvelebil M. 1986. Mesolithic prelude and neolithic revolution. In *Hunters in Transition: Mesolithic Societies of Temperate Eurasia and their Transition to Farming*, ed. M Zvelebil, pp. 5–15. Cambridge, UK: Cambridge Univ. Press
- Zvelebil M. 1989. On the transition to farming in Europe, or what was spreading with the Neolithic: a reply to Ammerman (1989). *Antiquity* 63:379–83
- Zvelebil M. 1993. Hunters or farmers? The Neolithic and Bronze Age societies of North-East Europe. In *Cultural Transformation and Interactions in Eastern Europe*, ed. J Chapman, P Dolukhanov, pp. 146–62. Aldershot, UK: Avebury
- Zvelebil M. 1995a. Farmers our ancestors and the identity of Europe. In *Cultural Identity and Archaeology*, ed. P Graves-Brown, S Jones, C Gamble, pp. 145–66. London: Routledge
- Zvelebil M. 1995b. Indo-European origins and the agricultural transition in Europe. In *Whither Archaeology?*, ed. M Kuna, N Venclova, pp. 173–203. Praha, Czech Republic: Inst. Archaeol.
- Zvelebil M. 1998a. Genetic and cultural diversity of Europe: a comment on Cavalli-Sforza. *J. Anthropol. Res.* 54:411–17

- Zvelebil M. 1998b. What's in a name: the Mesolithic, the Neolithic, and social change at the Mesolithic-Neolithic transition. See Edmonds & Richards 1998, pp. 1–36
- Zvelebil M. 2000. The social context of the agricultural transition in Europe. See Renfrew & Boyle 2000, pp. 57–79
- Zvelebil M, Zvelebil KV. 1988. Agricultural transitions and Indo-European dispersals. *Antiquity* 62:574–83





**Figure 6** Maximum parsimony tree for the major Y-chromosome haplogroups and their distribution in Europe. The West Eurasian parts of the tree are highlighted and color-coded. The haplogroup nomenclature of Underhill et al. (2000) is indicated beneath the tree; that of the Y Chromosome Consortium (2002) is within the frame; and that of Semino et al. (2000) is in parentheses. See Semino et al. (2000) for further details.