Genetic Evidence for Prehistoric Demographic Changes in Europe

Guido Barbujani

Department of Life Sciences and Biotechnologies, University of Ferrara, Ferrara, Italy

Abstract

Objectives: Two main models have been proposed to explain the origins of the patterns of genetic variation in Europe, one emphasizing Paleolithic and the other Neolithic immigration from the Southeast. In this paper, I summarize how the models developed and how they can help address some open questions. **Methods:** The rationale of the methods traditionally supporting the Neolithic and the Paleolithic models is discussed, and the evidence supporting either of them is reviewed. **Results:** Ancient DNA evidence proves for good that the studies traditionally supporting the Paleolithic model had serious methodological flaws. This does not imply that the alternative model is right, but rather calls for further analyses explicitly testing the two models against the genomic information now available. **Conclusions:** Questions that need to be addressed include whether the two main models differ enough to be discriminated by analyses of modern DNA diversity, and to what extent inferences from ancient mitochondrial DNA can be trusted in the absence of sufficient datasets of ancient nuclear DNA. The time seems ripe for the construction of a more complex (and hence more realistic) model, incorporating the possibility of different processes affecting different geographic locations at different times.

© 2014 S. Karger AG, Basel

Introduction

Geneticists have been debating the origins and evolution of the European population for more than three decades now, but the importance of this topic goes far beyond that. In fact, the study of the patterns of genetic variation in Europe is the first example of a large-scale research program crossing the disciplinary boundaries between biology and archaeology, and ultimately involving experts in such different fields as linguistics, anatomy, cultural anthropology, and demography. Genetics started contributing to this field in 1978, with the publication of the first synthetic maps of human allele frequencies [1]. A high degree of geographical structuring was apparent in Europe, with continent-wide clines which did not seem compatible with the simple effects of isolation by distance [2, 3].

But what were then the processes underlying those broad clinal patterns? Theory shows that demographic history affects all genome regions equally, whereas selection leaves specific signatures in a necessarily limited set of loci [4]. Therefore, only some kind of large-scale migration can generate nonrandom patterns at many loci, but how and especially when that happened was all but obvious. The initial consensus was that the shift from food collection to food production, or the Neolithic transition, starting in the Levant some 10,000 years BP, was the best candidate for a complex series of demographic changes, ultimately producing genetic gradients across much of the continent. Only after two decades, with the

Guido Barbujani
Department of Life Sciences and Biotechnologies
University of Ferrara
Via L. Borsari 46, IT–44100 Ferrara (Italy)
E-Mail g.barbujani@unife.it

appearance of the first datasets of mitochondrial DNA (mtDNA), an alternative proposal emerged; the European gradients might be older and reflect the first Paleolithic peopling of the continent, some 30,000 or 40,000 years BP. As a consequence, the spread of farming technologies across Europe in the Neolithic did not necessarily entail extensive gene flow and could be essentially regarded as a process of cultural transmission.

Of course, between an exclusively Paleolithic and an exclusively Neolithic origin of the patterns, there is a broad range of intermediate possibilities. However, the many migrational exchanges documented in the more recent archaeological and historical records do not seem relevant to understanding the origin of the clines, because none of them occurred on a sufficiently broad scale to determine genetic effects across all Europe [5]. In addition, historical migrations are documented in all directions [6] and so are unlikely to result in a coherent continental pattern, which, if anything, is there not *because of*, but *despite*, recent gene flow. Therefore, the basic alternative is between a mostly Neolithic and a mostly Paleolithic process, keeping in mind that reality is likely to be somewhere in between these necessary abstractions.

In this paper, I shall start by summarizing the rationale underlying the two main alternative views, here referred to as Neolithic and Paleolithic models [for a detailed description of the archaeological evidence on these processes, see Pinhasi et al. 7]. I shall then point out that the Paleolithic model, in its original formulation, suffered from fundamental theoretical flaws, which were immediately identified on theoretical grounds and have now been highlighted by ancient DNA evidence. I shall then move to an examination of other kinds of methodological problems, which emerged within the Neolithic model. Despite a personal preference for an interpretation of the gradients associating them to the effects of Neolithic gene flow, I will conclude that, at present, the exact weight of demographic changes occurring in Neolithic and Paleolithic times has not been firmly established. Finally, I shall list some questions which, in my opinion, need to be addressed if we are to proceed in this challenging but utterly stimulating area of investigation.

A Neolithic Demic Diffusion?

In 1978, Menozzi et al. [1] first summarized variation of many allele frequencies in Europe by principal component analysis (PCA). Starting from data on 38 independent alleles at 9 protein loci (ABO, Rh, MNS, Le, Fy, Hp, PGMi, HLA-A, and HLA-B), they identified a broad Southeast to Northwest cline spanning from the Near East to the Atlantic coasts and accounting for 27% of the total genetic variance. They also noticed a general correspondence between their PCA map and archaeological maps, showing the diffusion of farming activities from the Near East into Europe, and proposed that the latter explain the former. In other words, the structure evident in the genetic data might be a product of phenomena documented in the archaeological record, namely the Neolithic transition, i.e. the westward and northward spread of the artifacts testifying the shift from a food-collecting to a food-producing lifestyle.

In its essence, the idea was as follows. In principle, Neolithic technologies may have spread by cultural exchanges or by migration (most likely by a combination thereof, but still it is important to identify the main factor of spread); conversely, at the genetic level, the European clines could only be accounted for by a large-scale process of gene flow. Because cultural contacts alone cannot produce genetic gradients paralleling the diffusion of Neolithic artifacts, it was proposed that early Neolithic farmers brought their know-how, their genes [6], and perhaps their languages, too (see below), to Europe by demic diffusion.

In fact, what we shall refer to as the Neolithic model here is based on a complex mathematical framework, considering five main factors: (1) an origin of agriculture in the fertile crescent, some 10,000 years ago; (2) the existence of allele frequency differences at some loci between Near Eastern farmers and European hunter-gatherers; (3) a demographic growth in farming communities, prompted by the greater availability of resources; (4) a gradual dispersal of farmers towards the North and West, looking for new arable land, and (5) a lower rate of population increase for hunting-gathering than for farming people, even when, after the farmers' dispersal, they came to occupy the same territories (fig. 1a). In time, this process is expected (and has actually been shown) to generate continental clines [8, 9]. Then, after the demographic increase prompted by the possibility to produce food, densities were too high for successive migration waves to deeply affect the genetic buildup of the population [8]. Therefore, the spread of farming technologies in Europe was regarded as largely due to a demic, and not cultural, process; farming spread because farmers did.

For some time, a role of selection, with clines regarded as adaptations to variable environmental pressures, was also envisaged [10]. However, this seems a possible expla-
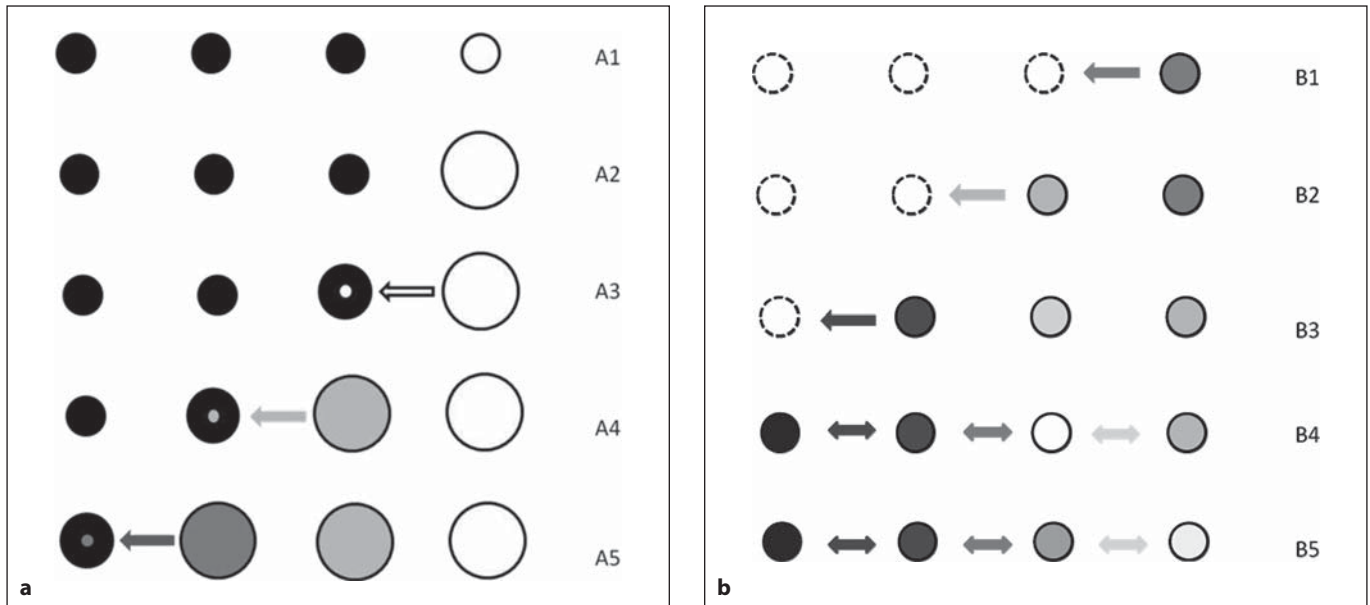
Fig. 1. Conditions for the origin of broad clines under Neolithic demic diffusion (**a**) and Paleolithic dispersal (**b**). Circles are populations, arranged along a geographic transect, and the circles' sizes are proportional to population sizes. The shades within circles represent the frequencies of an allele, with white = 0, black = 1. **a** The farming and hunting-gathering populations differ for allele frequencies; here, we have a fixed difference, but the model assumes any significant difference. At the beginning of the process (A1), all populations have the same size, but the Easternmost farmers use technologies allowing the territory to support a larger population. The farming population grows in size, until the carrying capacity of the area is reached (A2); farmers then disperse to the nearest locality occupied by hunter-gatherers, resulting in a moderate increase in the local population size (A3). However, because the two groups do not immediately admix, farmers retain a greater potential of demographic growth, and hence when admixture occurs, after a few generations, their alleles come to represent a larger share of the population's alleles (A4). The same happens again in successive diffusion steps (A5). **b** Here, dotted lines represent areas unoccupied by populations. Because of the small size of the expanding groups, each colonization step involves some degree of genetic drift. In the first colonization step (B1), there is a downward fluctuation of the allele frequency (B2), followed by an upward fluctuation (B3), while drift continues to affect all existing populations. When all suitable locations have been occupied, short-range gene flow (double arrows) begins (B4), ultimately leading to a clinal allele frequency distribution (B5).

nation for the variation observed at specific loci, such as that for lactase persistence in milk-drinking cultures [11]. At a more general level, parallel gradients at many independent loci strongly suggest that a common factor, hence not selection, shaped the Europeans' genome variation.

In the following years, several studies of allele frequencies gave further support to the Neolithic model [2, 8, 12–14], and independent pieces of evidence coming from other fields seemed to nicely fit this view. During the transition to a farming lifestyle, a fast increase in the proportion of immature skeletons is observed, in Europe as well as in other continents [15]. This finding is regarded as a signature of a fertility increase, ultimately resulting in an increase in population sizes [16], as expected under the original formulation of the Neolithic model

[8]. Also, a broad analysis of craniometric data showed an excellent fit with the predictions of a model of Neolithic demic diffusion [17]. The evidence for a westward and northward population dispersal appeared so convincing that it was proposed to reconsider linguistic evidence in the light of the genetic data, leading to what is now called the farming/language dispersal hypothesis [18, 19]. According to that hypothesis, one needs just a single population expansion to account for the diffusion of agricultural technologies, genes, and Indo-European languages from the Levant into Europe, a view recently corroborated by extensive analyses of linguistic diversity [20]. What was missing in this broad picture towards the end of the 20th century was direct evidence of European DNA variation.

Table 1. Estimated ages (×1,000 years) of the main European mtDNA haplogroups in four studies

Haplogroup	Ref. [21]	Ref. [22]	Ref. [23]	Ref. [24]
H	23.5 (1)	20.5	11.0–14.0	6.5–18.2
V	23.5 (1)	12.5	11.0–14.0	11.1–16.9
J	23.5 (2A)	8.0 (J*)	8.5	7.4–11.4
T	35.5 (2B)	46.5	11.0–14.0	10.9–19.2
IWX	50.5 (3)		11.0–14.0	
X		24.0		17.0–30.0
I		35.0	11.0–14.0	26.9–40.3
K	17.5 (4)	15.5	11.0–14.0	10.8–16.4
U	36.5 (5)	52.5	50.0	15.5–29.6

Haplogroup names in reference 21 are between parentheses, and so is a subgroup of haplogroup J (J*) in reference 22. In reference 24, haplogroup ages were estimated under four criteria for founder identification; figures in the last column refer to the criterion termed *f2*.

A Series of Paleolithic Founder Effects?

mtDNA was for many years the primary source of information on human molecular diversity and was essentially the only one available for comparing populations at the large geographical scale. Because mtDNA is basically unaffected by recombination, it is easy to summarize its diversity by evolutionary trees or networks and to estimate the age of mutations shared by the groups of variants that came to be called haplogroups. In the first analyses of mtDNA, very old ages were estimated for the main European haplogroups (table 1); these estimates somewhat fluctuated through time, reflecting changes in the database and in the definitions of the haplogroups, but few of them were <10,000 years old [21–24].

Richards et al. [24] then made two strong assumptions: (1) each cluster can be assigned, in its entirety, to one of the main migration phases (i.e. Neolithic and Paleolithic), and (2) the age of each cluster approximates very closely the timing of the migration event. Under those assumptions, often referred to as *phylogeographic approach*, these studies concluded that the main mitochondrial variants in Europe predate the Neolithic expansion, and hence that the genetic structure of Europe has been determined in Paleolithic times, with only haplogroup J and some subhaplogroups representing the effects of the Neolithic expansion. Since the frequency of these variants over Europe was 20%, this was also considered to be a measure of the fraction of the Europeans' ancestors who entered the continent in Neolithic times [24].

In these first studies, no significant spatial pattern was identified, and no attempt was made to reconcile the mtDNA results with previous data on protein diversity. Yet, leaving aside the validity of the assumptions of that kind of phylogeographic approach, to which we will come back, it is worth noting that the extensive geographic clines of nuclear markers are not incompatible with a Paleolithic model, in principle. Indeed, both theory and simulations agree that those patterns can be generated during the colonization of a new territory by small groups of individuals, i.e. through a series of founder effects. In this case, however, the role of gene flow would not be the same as under a Neolithic model; drift would first lead to fluctuations and fixations of alleles in the course of the range expansion, and successive gene flow would produce intermediate allele frequencies in intermediate locations [3] (fig. 1b). As a consequence, under a Paleolithic model, the spread of Neolithic technologies was attributed largely to cultural exchanges, a view also shared by influential archaeologists [25]. The H haplogroup, in particular, currently the most common European haplogroup, was proposed as the signature of the Paleolithic expansion in Europe [26].

These data also had an impact upon some of the original proposers of the Neolithic model. After expressing doubts on the reliability of mtDNA data for demographic reconstruction, motivated by the high mitochondrial mutation rate, Cavalli-Sforza and Minch [27] wrote that a figure close to 25% might realistically represent the contribution of Neolithic immigrants to the gene pool of Europeans, because, in PCA analyses of allele frequencies, the clinal component accounted for roughly one quarter of the genetic variance [1]. If so, the level of disagreement would have been minimal, if any; all such estimates come with a large standard error, and so the 20% proposed by Sykes [23] may not differ significantly from 25%. What was unclear in that line of reasoning was the relationship between PCA components and the origin of populations.

The first studies of nuclear DNA did not contribute much to clarifying the picture. Indeed, whereas broad clines were observed at several autosomal loci [28, 29], which were not suitable to be described by a phylogeographic approach, Y chromosome data were interpreted in contrasting manners and regarded as supporting either the Paleolithic [30, 31] or the Neolithic [32–34] model. Clearly, in the absence of formal tests of hypotheses, descriptions of modern genetic diversity were sufficient to generate a lively debate, but not to establish for good which phase of human prehistory had been most important in determining the current structure of the European population.

The Trouble with the Paleolithic Model

The conclusions of the first studies supporting the Paleolithic model rest crucially on the assumptions that each cluster of alleles can be entirely assigned to a single migration phase, and that its age approximately corresponds to the timing of the migration event [24]. But are these assumptions justified?

As a matter of fact, hardly so [35]. A mutation is a biochemical change affecting a fertile cell in an individual; a migration is a demographic process affecting many people. The two phenomena are independent; people of blood group A did not expand all over the world at the same time, right after mutation occurred, and that is also the case for mtDNA and Y chromosome mutations. Also, application of these principles to the data of table 1 leads to nonsense; some European mitochondrial haplogroups, and hence the entire mitochondrial genealogy, predate the arrival in Europe of the first anatomically modern humans (some 40,000 years ago), which can only mean that the relevant mutation events occurred out of Europe. Furthermore, the definition of haplogroups is arbitrary, depending on what one considers their founding mutation; by splitting a haplogroup in subhaplogroups of inferior order, clusters can be constructed with any lower time depth. Finally, the average coalescence time of two sequences sampled from two diverging populations is older, or much older, than the split of the groups [36, 37]. Unless a group colonizing a new territory passes through a strong and long-lasting bottleneck, part of its initial diversity will be maintained [38]. Therefore, the coalescence times inferred from samples of its descendants will be close to the coalescence times of the population of origin, consistently overestimating the age of the derived populations [36]. In short, people, not haplogroups, migrate, and hence inferences about population history must be based on measures of genetic diversity between populations, not between molecules.

Phylogeography, the study of the relationships between gene genealogies and geography [39], is informative on a number of evolutionary processes, but its assumptions do not hold when gene flow causes uncoupling of gene genealogies and population genealogies [40]; unfortunately, that is often the case with humans [35, 37]. The theoretical debate on these subjects has been going on for quite some time now, but ancient DNA evidence has demonstrated for good that the original assumptions of the Paleolithic model were wrong. The 105 Neolithic Europeans whose mtDNA molecule has been completely sequenced did not only carry the J haplogroup, but a broad set of haplotypes, including some of the very ancient haplogroups U and T; among 21 pre-Neolithic hunter-gatherers, there is no evidence of the mislabeled 'Paleolithic' haplogroup H, which seems instead to have become widespread in Europe only through Neolithic immigration [41]. Globally speaking, the genetic makeup of the contemporary populations is very similar to that of early Neolithic farmers and differs sharply from that of pre-Neolithic hunter-gatherers, both at the mtDNA level [41, 42, 66] and (for the only 6 samples analyzed so far) at the nuclear level [43, 44].

The Trouble with the Neolithic Model

The Paleolithic model was built on sand, but this does not imply that the alternative, Neolithic, model is in perfect shape. Already in 1999, Sokal et al. [45] had pointed out that the PCA synthetic maps inferred from allele frequencies require previous interpolation of the data; therefore, when PCA shows a gradient, it is generally unclear to what extent that gradient is really in the data or has been generated by interpolation. In addition, ordered patterns in PCA have recently been shown to occasionally arise under isolation by distance, and hence do not necessarily point to specific migration events [46]. Further, simulation work suggested that selection of polymorphisms with high frequencies of the rarer allele promotes the observation of genetic clines, which are not expected for random polymorphisms [47], casting doubts on the significance of Menozzi et al.'s [1] initial observations. Additional doubts on the Neolithic model derive from a recent analysis of the radiocarbon dates of Central and Northern European Neolithic sites, suggesting that population densities did not increase steadily with the advent of farming technologies, but had ample fluctuations [48]. The effects of such phenomena on long-term patterns of genetic variation have not been investigated yet, but these results, although not confirmed by all studies [48], call into question a core assumption of all Neolithic models, i.e. that the greater resource availability did immediately result in demographic growth.

As a consequence of all this, one has to conclude that, no matter how poorly the Paleolithic model was initially put forward, the Neolithic model has its problems, too. Clines do not come with a date, and so their existence, per se, does not prove that they arose in Neolithic times, nor even, in fact, by demic diffusion.

Comparing Models

Formal tests of the competing hypotheses based on explicit demographic modeling have been surprisingly rare so far. Using forward simulations of five evolutionary models in which only Indo-European speakers were considered, all incorporating isolation by distance, Barbujani et al. [3] showed that clines are best accounted for by two models where dispersal of Neolithic farmers from the Near East depends only on population growth. Models of greater complexity, where archaeological time data constrained the timing of the farmers' expansion, failed to explain a larger fraction of the observed genetic variation, and hence appear unnecessary. However, the same study showed that gradients can arise in two ways, i.e. not only by incomplete admixture between dispersing farmers and preexisting hunters and gatherers (as expected under Neolithic demic diffusion), but also by founder effects during a population expansion at low densities, hence not necessarily in Neolithic times [3] [see Liu et al. 49 for an example of how broad gradients in many indexes of genetic diversity may reflect a series of founder effects].

Currat and Excoffier [47] studied the expected distribution of coalescence times in Europe through a backward simulation based on the core algorithm of Barbujani et al. [3] and compared it with the distribution of mitochondrial coalescence times over Europe. They concluded that (a) a very small Paleolithic contribution to the original founder populations would be sufficient for most modern people to be descended from Paleolithic ancestors, and (b) the clinal distribution of allele frequencies might just reflect a bias in favor of highly polymorphic SNPs [47].

In the meantime, however, admixture coefficients had been estimated from autosomal, mitochondrial, and Y-linked polymorphisms, under a model regarding the European populations as derived from hybridization among 4 or less potential parental populations. Two main genome components became apparent, presumably corresponding to the contributions of the first, Paleolithic, Europeans, and of the early Neolithic farmers, the latter decreasing from East to West, as predicted by a model in which the alleles of Neolithic immigrants from the Near East got diluted during an expansion towards the Northwest [50].

It was only 8 years later that two further explicit tests were published, both based on comparisons of modern and ancient DNA in Iberian samples, and both indicating some level of demographic replacement associated with the arrival of farming technologies. Sánchez-Quinto et al. [44] typed >20,000 genomic SNPs in two 7,000-year-old Mesolithic individuals. Their statistical analyses, based on approximate Bayesian computation, showed that a model in which Neolithic farmers have a greater demographic growth rate and largely replace preexisting Paleolithic people is >3,000 times more likely than a model of genetic continuity in Europe since Paleolithic times. Gamba et al. [51] worked on mtDNA from 13 Neolithic individuals and found greater support for a model involving a higher rate of demographic growth for Near Eastern Neolithic immigrants than for two alternative models assuming either panmixia or admixture between two groups with equal rates of demographic growth.

Clearly, it is difficult to draw any firm conclusions from these contrasting results. The sharp differences between forward and backward simulations based on similar modeling of the demographic processes are particularly puzzling, and it may be argued that factors considered secondary in the modeling had instead a large effect on the results. One such factor may be the exact rate of acculturation of hunter-gatherers, turning them into farmers and thus giving them an increased potential of demographic growth. At any rate, older expansions tend to conceal the effects of more recent demographic changes [52], and so the evidence for Paleolithic expansions [47] is not incompatible with a further expansion occurring in the Neolithic period.

What Next?

Many papers end up with the disappointing statement that we need more data. Still, nobody can deny that a larger autosomal DNA dataset covering much of Europe, in time and place, would help us understand much better the European demographic history. Alas, it will take time to assemble it, and while we wait, we must try to make sense of the limited existing data. There is little doubt that Neolithic processes had a major impact on the genetic structure of Europeans, but these processes have not had the same impact upon all geographical regions; in addition, the resulting structure was not fixed then forever, but kept being reshaped by subsequent, smaller-scale, phenomena of gene flow and drift [7]. What follows, then, is a list of questions I deem relevant, some of them already addressed, others, in my opinion, requiring further attention.

One question concerns the limitations of PCA analysis. Above and beyond the effects of interpolation of allele frequencies, we have seen that ordered patterns in PCA

Barbujani

may also arise by chance, and hence do not necessarily reflect specific migration events [46]. In addition, a bias in favor of the most polymorphic markers may give the impression of clinal variation, even when the majority of genomic markers do not show any pattern [47]. Yet, the existence of broad genetic gradients across much of Europe has been confirmed by different methods [53, 54] and using polymorphisms such as microsatellites [32], the mutation rates of which seem high enough to make the ascertainment bias irrelevant for demographic reconstructions [55]. Thus, no matter how shaky the PCA results are considered, the clines are out there and call for an explanation.

Another important question is whether a single process accounting for the European genetic structure is too much of an oversimplification, and here the answer seems yes. Three mitochondrial studies, based on ancient [56, 57] and modern [58] DNA, found evidence for more complex Neolithic processes. Apparently, demic diffusion prevailed in Southern Europe, while in Central Europe cultural transmission may have been the rule. However, extensive genomic data from 5,000-year-old Scandinavian hunter-gatherers and farmers showed that they fall genetically close to modern Northern and modern Southern Europeans, respectively, suggesting that agriculture was introduced in Scandinavia by people immigrating from the South [43]. Incidentally, this may imply that results of several studies may have shifted in favor of either model, depending on the relative weight of Southern, Central, and Northern Europeans among the studied samples. Nongenetic evidence suggests that the causes of this complex pattern are also complex. First, zones of geographical heterogeneity are known to cause resistance to demic diffusion [59]; secondly, archaeological data suggest multiple waves of Neolithic expansion [60], and finally, rates of farming expansions appear inversely correlated with the intensity of preexisting agricultural activities [61]. All this led Fort [62] to conclude that although cultural diffusion of farming technologies did actually occur, demic diffusion was the most important mechanism in this historical process at the continental scale. The open question, now, is what exactly happened, and in which areas of Europe the transition to farming was not caused by immigration (as was probably the rule) but by a technological shift in the previously hunting and gathering population.

Therefore, moving to less schematic models currently seems the main research priority in this area. This means asking in the first place whether the two main models are different enough to be discriminated by analyses of mod-

ern DNA diversity. In other words, because both Paleolithic and Neolithic migration occurred largely along the same Southeast to Northwest axis and samples are scattered over a broad space, it is all but clear whether the available methods and dataset can safely tell us which of the main models is better. It seems possible to address this question by estimating the type 1 error [63], i.e. generating by simulation pseudo-datasets under the two models, and then testing how often the generating model can be inferred from a blind analysis of these data [see an example in Zeder 60].

While we wait for sufficiently large datasets of ancient nuclear DNA to be assembled, how much can we trust inferences from ancient mtDNA? This is a third open question. True, mtDNA is transmitted as a single locus, its diversity represents but one realization of an evolutionary process with a high stochastic component, and so its analysis may or may not be very informative on the whole population's history [64, 65]. Yet, it will take long until autosomal polymorphisms are mapped at a comparable geographical detail, and in the meantime, one locus is better than no loci at all. There currently seems to be no alternative to taking seriously mtDNA evidence, leaving open the possibility that the demographic history of European males and females might have been somewhat different [see, however, the study by Rasteiro and Chikhi 34, where a remarkable parallelism was found in the admixture histories inferred from mtDNA and Y chromosome data].

Fourth, computer simulations are indispensable to actually test hypotheses, as opposed to finding stories that may explain the data, but one should be aware that spatial simulations come in different flavors, each with its pros and cons. Forward simulations [3, 11, 34], from the past to the present, are time consuming and complicated to handle, and hence it is often convenient to turn to coalescent-based simulations, in which the fate of a sample of DNA fragments is followed backwards from the present time to its common ancestor. However, this approach, in many ways reliable and certainly faster than its alternative, has shortcomings too. One is that the spatial model is generally simulated once, and then many coalescent simulations are run based on the results of that single simulation. This way, the stochasticity of the genealogical process is taken into account, but stochastic effects in the course of the spatial process are disregarded; it is all but clear how deeply this can bias the simulations' results. In addition, as we have seen, information that is often difficult or impossible to infer from archaeological evidence (was there a reproductive barrier between farmers and

hunter-gatherers when the two populations met, and how long did it last?) may turn out to deeply influence the simulation results [compare Barbujani et al. 3 and Currat and Excoffier 47].

Keeping in mind both the potentials and limitations of current data, it seems about time for geneticists, archaeologists, and paleodemographers to develop a mixed model, incorporating the possibility of different population processes affecting different geographic locations in Paleolithic and Neolithic times. It will then be a matter of years until enough genomic data are assembled, and one can move to testing whether this mixed model is able to account for the existing patterns better than alternative models. The original Neolithic and Paleolithic models were very useful for many years to interpret the data and plan research, but it seems now both necessary and possible to refine them.

Acknowledgments

References

1 Menozzi P, Piazza A, Cavalli-Sforza L: Synthetic maps of human gene frequencies in Europeans. Science 1978;201:786–792.
2 Sokal RR, Harding RM, Oden NL: Spatial patterns of human gene frequencies in Europe. Am J Phys Anthropol 1989;80:267–294.
3 Barbujani G, Sokal RR, Oden NL: Indo-European origins: a computer-simulation test of five hypotheses. Am J Phys Anthropol 1995;96:109–132.
4 Harris EE, Meyer D: The molecular signature of selection underlying human adaptations. Am J Phys Anthropol 2006;suppl 43:89–130.
5 Sokal RR, Oden NL, Walker J, Di Giovanni D, Thomson BA: Historical population movements in Europe influence genetic relationships in modern samples. Hum Biol 1996;68:873–898.
6 Sokal RR: Ancient movement patterns determine modern genetic variances in Europe. Hum Biol 1991;63:589–606.
7 Pinhasi R, Thomas MG, Hofreiter M, Currat M, Burger J: The genetic history of Europeans. Trends Genet 2012;28:496–505.
8 Ammerman AJ, Cavalli-Sforza LL: The Neolithic Transition and the Genetics of Populations in Europe. Princeton, Princeton University Press, 1984.
9 Ammerman AJ, Cavalli-Sforza LL: A population model for the diffusion of early farming in Europe; in Renfrew C (ed): The Explanation of Culture Change. London, Duckworth, 1973, pp 345–357.
10 Fix AG: Gene frequency clines in Europe: demic diffusion or natural selection? J Royal Anthropol Inst 1996;2:625–643.
11 Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG: The origins of lactase persistence in Europe. PLoS Comput Biol 2009;5:e1000491.
12 Barbujani G, Pilastro A, De Domenico S, Renfrew C: Genetic variation in North Africa and Eurasia: Neolithic demic diffusion vs Paleolithic colonisation. Am J Phys Anthropol 1994;95:137–154.

13 Sokal RR, Menozzi P: Spatial autocorrelation of HLA frequencies in Europe supports demic diffusion of early farmers. Am Nat 1982;119:1–17.
14 Cavalli-Sforza LL, Menozzi P, Piazza A: Demic expansions and human evolution. Science 1993;259:639–646.
15 Bellwood P, Oxenham M: The expansions of farming societies and the role of the Neolithic demographic transition; in Bocquet-Appel JP, Bar-Yosef O (eds): The Neolithic Demographic Transition and Its Consequences. Berlin, Springer, 2008, pp 13–34.
16 Bocquet-Appel JP: The demographic impact of the agricultural system in human history. Curr Anthropol 2009;50:657–660.
17 Pinhasi R, von Cramon-Taubadel N: Craniometric data supports demic diffusion model for the spread of agriculture into Europe. PLoS One 2009;4:e6747.
18 Renfrew C: Archaeology and Language: The Puzzle of Indo-European Origins. London, Jonathan Cape, 1987.
19 Bellwood PS, Renfrew C: Examining the Farming/Language Dispersal Hypothesis. Cambridge, McDonald Institute for Archaeological Research, 2002.
20 Bouckaert R, Lemey P, Dunn M, Greenhill SJ, Alekseyenko AV, Drummond AJ, Gray RD, Suchard MA, Atkinson QD: Mapping the origins and expansion of the Indo-European language family. Science 2012;337:957–960.
21 Richards M, Corte-Real H, Forster P, Macaulay VA, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt HJ, Sykes B: Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 1996;59:185–203.
22 Richards M, Macaulay VA, Bandelt HJ, Sykes BC: Phylogeography of mitochondrial DNA in Western Europe. Ann Hum Genet 1998;62:241–260.

23 Sykes BC: The molecular genetics of European ancestry. Phil Trans Royal Soc London B 1999;354:131–139.
24 Richards M, Macaulay VA, Hickey E, et al: Tracing European founder lineages in the Near East mtDNA pool. Am J Hum Genet 2000;67:1251–1276.
25 Zvelebil M: On the transition to farming in Europe, or what was spreading with the Neolithic: a reply to Ammerman. Antiquity 1989;63:379–383.
26 Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon JM, Roostalu U, Loogväli EL, Kivisild T, Bandelt HJ, Richards M, Villems R, Santachiara-Benerecetti AS, Semino O, Torroni A: The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am J Hum Genet 2004;75:910–918.
27 Cavalli-Sforza LL, Minch E: Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 1997;61:247–251.
28 Chikhi L, Destro-Bisol G, Bertorelle G, Pascali V, Barbujani G: Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. Proc Natl Acad Sci USA 1998;95:9053–9058.
29 Belle EM, Landry PA, Barbujani G: Origins and evolution of the Europeans' genome: evidence from multiple microsatellite loci. Proc R Soc B 2006;273:1595–1602.
30 Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA: The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 2000;290:1155–1159.

31 Battaglia V, Fornarino S, Al-Zahery N, Olivieri A, Pala M, Myres NM, King RJ, Rootsi S, Marjanovic D, Primorac D, Hadziselimovic R, Vidovic S, Drobnic K, Durmishi N, Torroni A, Santachiara-Benerecetti AS, Underhill PA, Semino O: Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. Eur J Hum Genet 2009;17: 820–830.

32 Chikhi L, Nichols RA, Barbujani G, Beaumont MA: Y genetic data support the Neolithic demic diffusion model. Proc Natl Acad Sci USA 2012;99:10008–10013.

33 Balaresque P, Bowden GR, Adams SM, Leung HY, King TE, Rosser ZH, Goodwin J, Moisan JP, Richard C, Millward A, Demaine AG, Barbujani G, Previderè C, Wilson IJ, Tyler-Smith C, Jobling MA: A predominantly Neolithic origin for European paternal lineages. PLoS Biology 2010;19:e1000285.

34 Rasteiro R, Chikhi L: Female and male perspectives on the Neolithic transition in Europe: clues from ancient and modern genetic data. PLoS One 2013;8:e60944.

35 Goldstein DB, Chikhi L: Human migrations and population structure: what we know and why it matters. Annu Rev Genomics Hum Genet 2002;3:129–152.

36 Barbujani G, Bertorelle G, Chikhi L: Evidence for Paleolithic and Neolithic gene flow in Europe. Am J Hum Genet 1998;62:488–491.

37 Barbujani G, Goldstein DB: Africans and Asians abroad: genetic diversity in Europe. Annu Rev Genom Human Genet 2004;5:119–150.

38 Nichols RA: Gene trees and species trees are not the same. Trends Ecol Evol 2001;16:358–364.

39 Avise J: Phylogeography: The History and Formation of Species. Cambridge, Harvard University Press, 2000.

40 Nielsen R, Beaumont MA: Statistical inferences in phylogeography. Mol Ecol 2009;18: 1034–1047.

41 Fu Q, Rudan P, Pääbo S, Krause J: Complete mitochondrial genomes reveal Neolithic expansion into Europe. PLoS One 2012;7: e32473.

42 Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind CJ, Lueth F, Terberger T, Hiller J, Matsumura S, Forster P, Burger J: Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. Science 2009;326:137–140.

43 Skoglund P, Malmström H, Raghavan M, Stora J, Hall P, Willerslev E, Gilbert MTP, Götherström A, Jakobsson M: Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science 2012;336:466–469.

44 Sánchez-Quinto F, Schroeder H, Ramirez O, Ávila-Arcos MC, Pybus M, Olalde I, Velazquez AM, Prada Marcos ME, Vidal Encinas JM, Bertranpetit J, Orlando L, Gilbert MT, Lalueza-Fox C: Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. Curr Biol 2012;22:1494–1499.

45 Sokal RR, Oden NL, Thomson BA: A problem with synthetic maps. Hum Biol 1999;71:1–13.

46 Novembre J, Stephens M: Interpreting principal component analysis of spatial population genetic variation. Nature Genet 2008;40:646–649.

47 Currat M, Excoffier L: The effect of the Neolithic expansion on European molecular diversity. Proc R Soc B 2005;272:679–688.

48 Shennan S, Downey SS, Timpson A, Edinborough K, Colledge S, Kerig T, Manning K, Thomas MG: Regional population collapse followed initial agriculture booms in mid-Holocene Europe. Nat Commun 2013;4:2486.

49 Liu H, Prugnolle F, Manica A, Balloux F: A geographically explicit genetic model of worldwide human-settlement history. Am J Hum Genet 2006;79:230–237.

50 Dupanloup I, Bertorelle G, Chikhi L, Barbujani G: Estimating the impact of prehistoric admixture on the Europeans' genome. Mol Biol Evol 2004;21:1361–1372.

51 Gamba C, Fernández E, Tirado M, Deguilloux MF, Pemonge MH, Utrilla P, Edo M, Molist M, Rasteiro R, Chikhi L, Arroyo-Pardo E: Ancient DNA from an early Neolithic Iberian population supports a pioneer colonization by first farmers. Mol Ecol 2012;21:45–56.

52 Rogers AR, Harpending HC: Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 1992;9:552–569.

53 Sokal RR, Oden NL, Wilson C: Genetic evidence for the spread of agriculture in Europe by demic diffusion. Nature 1991;351:143–145.

54 Lao O, Lu TT, Nothnagel M, et al: Correlation between genetic and geographic structure in Europe. Curr Biol 2008;18:1241–1248.

55 Romero IG, Manica A, Goudet J, Handley LL, Balloux F: How accurate is the current picture of human genetic variation? Heredity 2009; 102:120–126.

56 Sampietro ML, Lao O, Caramelli D, Lari M, Pou R, Martí M, Bertranpetit J, Lalueza-Fox C: Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. Proc Biol Sci 2007;274:2161–2167.

57 Hervella M, Izagirre N, Alonso S, Fregel R, Alonso A, Cabrera VM, de la Rua C: Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the Neolithic expansion into Europe. PLoS One 2012; 7:e34417.

58 Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G: Geographic patterns of mtDNA diversity in Europe. Am J Hum Genet 2000;66:262–278.

59 Ackland GJ, Signitzer M, Stratford K, Cohen MH: Cultural hitchhiking on the wave of advance of beneficial technologies. Proc Natl Acad Sci USA 2007;104:8714–8719.

60 Zeder MA: Domestication and early agriculture in the Mediterranean basin: origins, diffusion, and impact. Proc Natl Acad Sci USA 2008;105:11597–11604.

61 Bocquet-Appel JP, Naji S, Vander Linden M, Kozlowski J: Understanding the rates of expansion of the farming system in Europe. J Archaeol Sci 2011;39:531–546.

62 Fort J: Synthesis between demic and cultural diffusion in the Neolithic transition in Europe. Proc Natl Acad Sci USA 2012;109: 18669–18673.

63 Ghirotto S, Tassi F, Fumagalli E, Colonna V, Sandionigi A, Lari M, Vai S, Petiti E, Corti G, Rizzi E, De Bellis G, Caramelli D, Barbujani G: Origins and evolution of the Etruscans' mtDNA. PLoS One 2013;8:e55519.

64 Harpending H: Anthropological genetics: present and future; in Crawford MH (ed): Anthropological Genetics. Theory, Methods and Applications. Cambridge, Cambridge University Press, 2007, pp 456–466.

65 Balloux F: The worm in the fruit of the mitochondrial DNA tree. Heredity 2009;104:419–420.

66 Brandt G, Haak W, Adler CJ, Roth C, Szécsényi-Nagy A, Karimnia S, Möller-Rieker S, Meller H, Ganslmeier R, Friederich S, Dresely V, Nicklisch N, Pickrell JK, Sirocko F, Reich D, Cooper A, Alt KW, Genographic Consortium: Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. Science 2013;342:257–261.