Review TRENDS in Genetics Vol.22 No.6 June 2006

Harvesting the fruit of the human mtDNA tree

Antonio Torroni[1], Alessandro Achilli[1], Vincent Macaulay[2], Martin Richards[3] and Hans-Jürgen Bandelt[4]

[1]Dipartimento di Genetica e Microbiologia, Università di Pavia, Via Ferrata 1, 27100 Pavia, Italy
[2]Department of Statistics, University of Glasgow, University Avenue, Glasgow, UK, G12 8QQ
[3]Institute of Integrative and Comparative Biology, University of Leeds, L.C. Miall Building, Leeds, UK, LS2 9JT
[4]Department of Mathematics, University of Hamburg, Bundesstr.55, 20146 Hamburg, Germany

Human mitochondrial DNA (mtDNA) studies have entered a new phase since the blossoming of complete genome analyses. Sequencing complete mtDNAs is more expensive and more labour intensive than restriction analysis or simply sequencing the control region of the molecule. But the efforts are paying off, as the phylogenetic resolution of the mtDNA tree has been greatly improved, and, in turn, phylogeographic interpretations can be given correspondingly greater precision in terms of the timing and direction of human dispersals. Therefore, despite mtDNA being only a fraction of our total genome, the deciphering of its evolution is profoundly changing our perception about how modern humans spread across our planet. Here we illustrate the phylogeographic approach with two case studies: the initial dispersal out of Africa, and the colonization of Europe.

Introduction

In human cells, most genes (~25 000) are confined to the nucleus, limited to two copies per cell and transmitted equally from both parents according to Mendel's laws. A major exception is represented by the 37 mitochondrial DNA (mtDNA) genes. They are located within the mitochondria, organized in a small (~16.6 Kb) circular molecule of DNA, which is present in hundreds to thousands of copies per cell, and transmitted as a non-recombining unit only through the mother. Furthermore, human mtDNA is characterized by a much greater evolutionary rate than that of the average nuclear gene. Thus, its sequence variation has been generated solely by the sequential accumulation of new mutations along radiating maternal lineages. The time frame of these mutations is such that human mtDNA contains a molecular record not only of the genealogical history but also of the migrations of women who transmitted it through the generations. Because the process of molecular differentiation is relatively fast and occurred mainly during and after the recent process of dispersal into different parts of the world, subsets of mtDNA variation usually tend to be restricted to particular geographic areas and populations. Recent analyses of this variation at the highest possible level of molecular resolution (i.e. that of complete mtDNA sequences) are now enabling us to determine where and when particular branching events are likely to have occurred, as the first step in the reconstruction of prehistoric human dispersals.

Corresponding author: Torroni, A. (torroni@ipvgen.unipv.it).
Available online 4 May 2006

Human mtDNA variation

The earliest mitochondrial DNA research began by digesting the entire molecule with just a few restriction enzymes. As early as 1980, Brown's pioneering study of 21 humans from diverse ethnic and geographic backgrounds indicated that mtDNA restriction-enzyme fragment length polymorphism (RFLP) patterns could be used to trace human genetic history [1]. On the basis of the observed diversity in a worldwide sample, he obtained a surprisingly recent coalescence age estimate of ~180 000 years for global mtDNA variation.

During the early 1990s, in parallel with the studies addressing the origin of *Homo sapiens*, mtDNA analyses also began to be applied to individual continents on large numbers of subjects, with the objective of determining human origins in each major geographic area. These studies were performed by using either RFLP survey of the entire mtDNA molecule, thus mainly targeting the coding region, or sequencing of the first hypervariable segment (HVS-I) of the non-coding and fast-evolving control region. The rich variation seen in this short (~350 bp) segment, however, is accompanied by high levels of recurrent mutation, blurring the structure of the tree – while there is insufficient variation to distinguish many important ancient branches within the tree. Consequently, the phylogeographic signal of mtDNA as seen through the window of HVS-I variation can be almost invisible. A combination of high-resolution RFLP screening with HVS-I sequencing eventually resulted in a much more refined picture of the mtDNA world phylogeny – one in which the major branches of the tree were usually restricted geographically, some to sub-Saharan Africans, others to East Asians and yet others to Europeans and Near Easterners [2].

The first large-scale population studies were performed in Native Americans and focused on the origin, timing and numbers of ancestral migrations from Asia [3]. It was in this specific context that the (by now) universally accepted

 0168-9525/$ - see front matter doi:10.1016/j.tig.2006.04.001

mtDNA haplogroup nomenclature was initiated, with four basal (or primary) branches in the tree, named alphabetically as A, B, C and D. (A haplogroup is a group of mtDNA haplotypes derived by descent from the same ancestral mtDNA molecule as revealed by the sharing of a distinguishing mutational motif.) Later, the haplogroup structures of other continental populations were characterized [4,5] and cladistic rules for the hierarchical ordering of haplogroups and sub-haplogroups were explicitly established [6]. The same principles were later used as the basis of the reformed nomenclature for the Y chromosomal tree [7]. The naming of new haplogroups has since then naturally developed in a self-organizing way, ushered forward by those who produce novel data, following the rules of the nomenclature system and respecting the published record.

In the medical field, a team led by Ozawa had begun to resolve the Japanese mtDNA tree based on complete mtDNA sequencing [8–10], which eventually culminated in the recent study of Tanaka *et al.* [11]. However, that early work was almost unnoticed by the population genetics community until recently [12,13] when complete genome analyses started blossoming [14–16]. There are now >2000 complete mtDNA sequences published, so that the basal branching structure of mtDNA variation in many – perhaps most – parts of the world is now well understood. Details concerning specific branches of the phylogeny, often restricted to microgeographic contexts, are available from several recent studies [17–26]. As a result, not only the Out-of-Africa debate but also the controversy about the origins of Europeans has come into sharper focus, so that opposing bold claims of the past can give way to more refined modelling and insights, as we will outline next.

Archaeogenetics

Amorim [27] coined the elegant term 'archaeogenetics' for 'the newly-emerged discipline that applies molecular genetics to the study of the human past' [28], thus substituting the earlier term 'historical genetics' [29]. Archaeogenetics was prefigured in the work of Luca Cavalli-Sforza and his colleagues, working on classical genetic markers in the 1970s (see Ref. [30] for a personal account; and Ref. [31] for a summary of the classical work). Cavalli-Sforza's work was the first major attempt to bring archaeology and genetics together in a single discipline and, perhaps most significantly, he did this by forging collaborations with the archaeological community – an achievement of which many subsequent researchers in the field unfortunately failed to take note.

Cavalli-Sforza and his colleagues focused particularly on the settlement of Europe, using the methodology of synthetic genetic maps: geographic maps of isopleths of principal components (PCs) values of variation. Although ideally suited to the kind of data available at the time, this approach could not provide a time frame for the different components, so that the interpretation of the maps remained ambiguous [28,32,33]. The map of the first PC, accounting for ~27% of the total variation in classical marker frequencies across Europe and the Near East, has become something of an icon in the study of European origins. It showed a gradient from the southeast to the northwest, with the Near East at one pole and Europe at the other, which was originally seen as strong evidence for the 'demic diffusion' hypothesis – implying that Neolithic farmers coming from the Near East and autochthonous Mesolithic European hunter-gatherers intermingled progressively. We will return to this issue later, but first we move to the other classic issue in archaeogenetics: the exit out of Africa of anatomically modern humans.

Out of Africa

Travelling up the human mtDNA phylogeny from the root, corresponding to what has, often rather confusingly, been identified with 'mitochondrial Eve' [34], one passes through several bi- and trifurcations until one reaches the first multifurcation node, the root of haplogroup L3 (Figure 1). This root gave rise to many (successful) descendant haplogroups, perhaps reflecting some colonization event or local population growth ~80 000–90 000 years ago, which might have been triggered by the glacial interstadial phase 21 [35,36] and, possibly, followed by cultural adaptation in response to the changing climate. This event probably occurred in the Horn of Africa because the richest branching around the L3 node is reflected in modern Ethiopian samples [37], in contrast to the somewhat narrower mtDNA pools of other parts of sub-Saharan Africa.

Two splinters of this basal L3 variation, haplogroups M and N, cover the mtDNA pool of all non-Africans (excluding the descendants of migrations from Africa within the past few thousand years, such as those engendered by various slave trades). Haplogroup N almost immediately diverged further to give rise to haplogroup R (Figure 2). All over Eurasia, America, Australasia and Oceania descendants of all three corresponding root haplotypes (M, N and R) can be found, with only a few more impoverished regional exceptions. Therefore, these three root types (which originated somewhere between East Africa and the Persian Gulf) must have been the founder types for the Eurasian settlement ~60 000–65 000 years ago [22,38]. Not only are the descendants of these founders ubiquitous outside Africa but also their numerous basal sub-branches are more or less specific to the major geographical regions. There is now growing evidence that the richest basal variation in the three founder haplogroups M, N and R is found along the southern stretch of Eurasia, particularly in the Indian subcontinent [18,25]. Although Southeast Asia has been less comprehensively analysed to date, recent first results indicate a similarly high basal diversification in this region [22]. This suggests a rapid colonization along the southern coast of Asia, reaching Sahul ~60 000 years ago. The expansions northwards to fill the heartland of the continent occurred later, ~45 000 years ago, when technology and climatic conditions enabled the exploration of the interior of Eurasia. One of the more marginal extensions eventually led to the peopling of Europe.

Into Europe

Early mtDNA studies of European populations revealed that all Europeans essentially share the same set of

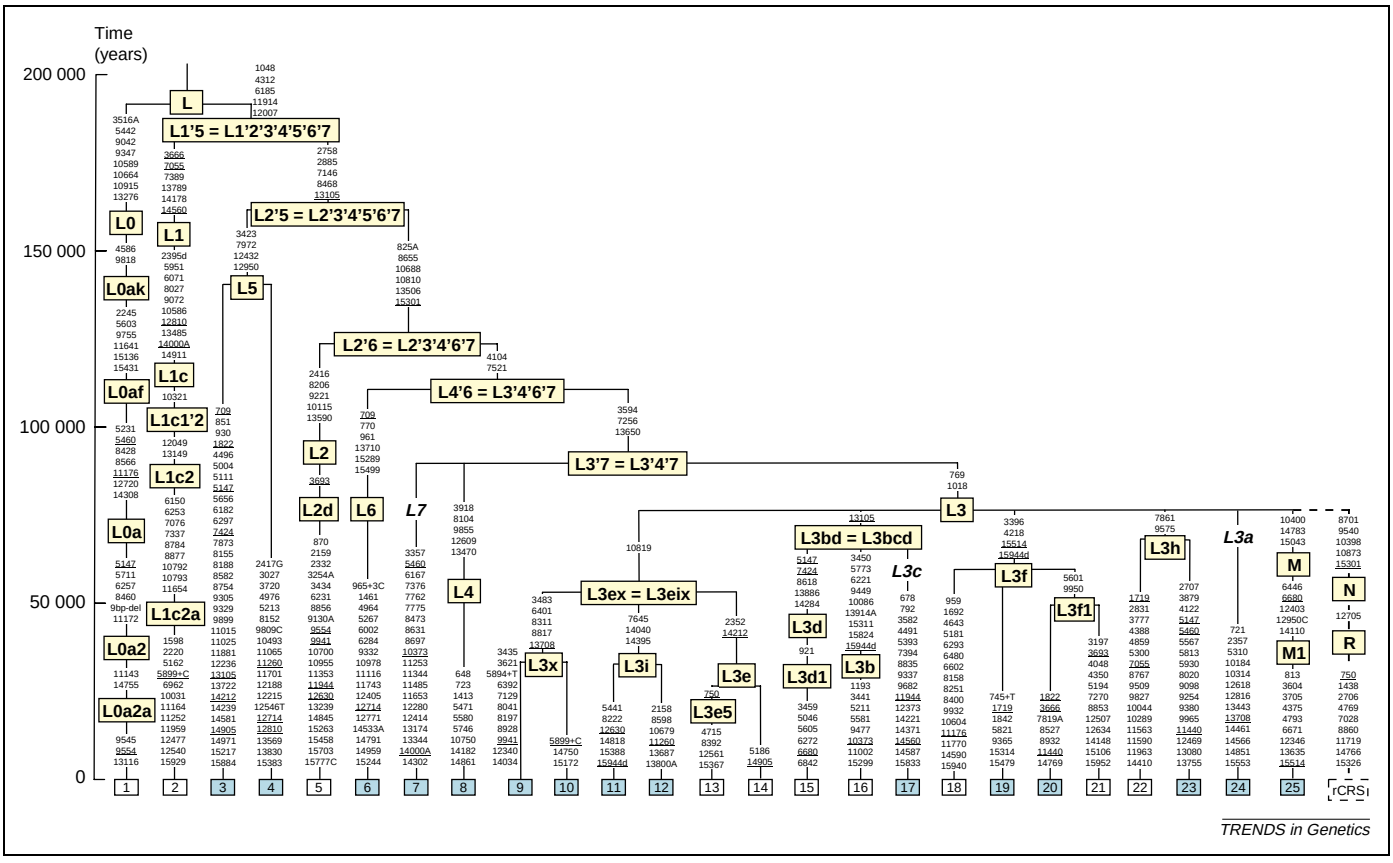

Figure 1. A maximum-likelihood phylogeny of African mtDNA haplogroups, with a calibrated time-scale (left). Sequencing of the 25 entire mtDNAs (1–25) was performed as in [17,20,22]. To construct the phylogeny, we focused on the more slowly evolving coding region of the molecule between nucleotide positions (nps) 577 and 16023, using two chimpanzee sequences to root the whole tree as described in Ref. [22]. A most parsimonious reconstruction of the coding-region mutations is shown on the branches. Mutations are scored relative to the revised Cambridge Reference Sequence (rCRS) [63] and therefore variant nucleotides do not necessarily constitute derived states; the rCRS is shown by the dashed lines. Suffixes indicate transversions (to A, G, C or T) or indels (+, d); recurrent mutations are underlined. The branching nodes in the phylogeny are shown in yellow, with the mutations above the node being shared by the derived haplogroups. The naming of haplogroups follows the scheme of [6], opting for the most compact notation. L3a and L3c denoting novel L3 branches from this study should not to be confused with the obsolete definitions of 'L3a' and 'L3c' (=U6a) introduced by Ref. [64]. The ancestral motifs for the novel haplogroups L7 (corresponding to L4g in Ref. [37]), L3a, and L3c (in italics) have not been determined. However, note that L4 and L7 (L4g) were clustered together in Ref. [37] because of their sharing of the (unstable) control-region mutation 16362. In addition, observe that there are minor differences with Ref. [21] in characterizing the mutations for some haplogroups. MtDNA sequences in sky-blue colour boxes are from Ethiopian subjects, whereas those in white boxes are from Nigerian (sequences 13, 18 and 21) and Dominican (sequences 1–2, 5, 14–16 and 22) subjects. Complete mtDNA sequences, including their control regions, are available in GenBank, accession numbers DQ341058–DQ341082.

haplogroups as Near Easterners and that those haplogroups were absent from sub-Saharan Africans and East Asians [39]. These observations confirmed that Europeans and Near Easterners share a rather recent common ancestry. From the complete mtDNA phylogeny that is now available, we can see that European mtDNA variation – embedded as it is in the West Eurasian mtDNA variation – is surprisingly impoverished in the number of independent basal lineages, compared with the South Asian mtDNA variation. Most European mtDNA lineages stem from just three R branches, haplogroups R0 (previously known as pre-HV), JT and U, in addition to three minor N branches, N1, N2 and X (Figure 2). Some further West Eurasian branches, such as R1 and R2 [18], seem to be rare in Europe. This reduction in basal variation can be best explained by the peripheral role that the pioneer migration into the Near East and subsequently Europe must have had in the broader 'Out-of-Africa' scenario (in contrast to traditional accounts that gave the Near East a pivotal role). The variation in the basal European mtDNA haplogroups can be dated to ~40 000–50 000 years, likely to be coincident with interstadial 12 [35,36].

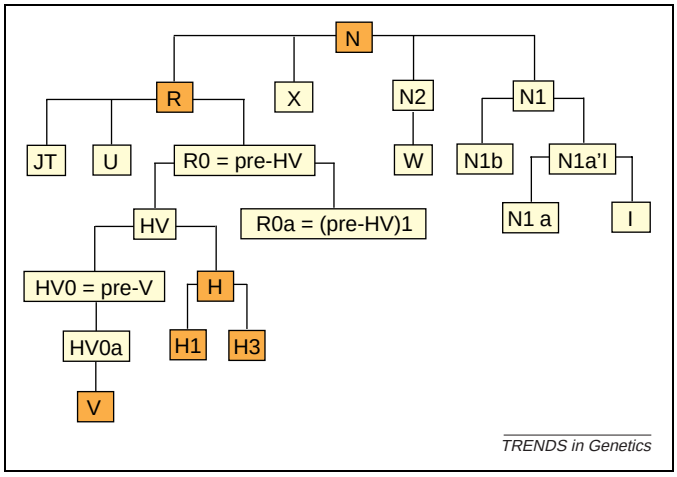

Figure 2. Schematic phylogeny of basal European mtDNA haplogroups. The haplogroups are denoted as in Ref. [18], but with new names that replace those beginning with 'pre-'. Nodes with a high degree of multifurcation (as testified by at least six emanating independent branches) are shaded in orange. In addition to H1 and H3 (see main text), other H sub-groups [17] could represent multifurcation nodes. However, sufficient complete mtDNA sequence data are not yet available to define the issue.

This, however, does not yet clarify to what extent the first migration of modern humans into Europe shaped the mtDNA pool of modern Europeans, because later immigration from the Near East could have replaced the descendants of the first settlers.

Some time after the demic diffusion hypothesis had been proposed, a debate arose on the issue of the relative contributions of Neolithic versus Palaeolithic gene pools in modern Europeans. What was the genetic contribution of the incipient Neolithic people, who gradually developed proto-forms of agriculture in the Near East ~11 000 years ago, relative to that of the indigenous European hunter-gatherers, when the Neolithic package first spread into Europe ~8000 years ago? This debate – fuelled initially by mtDNA and later by Y chromosome data – has become acrimonious at times. The received wisdom of the nineties, positing that the European gene pool predominantly resulted from Near Eastern newcomers bringing the Neolithic culture with them, was first challenged on the basis of mitochondrial HVS-I variation [40]. The new suggestion was that only a minority of lineages were a result of the Neolithic immigration, whereas the remaining lineages – dating back to between 15 000 and 50 000 years ago – seemed to have autochthonous roots of Early, Middle or Late Upper Palaeolithic origin within Europe. The results were, however, rather tentative because they relied on comparisons with a small and inadequate sample from the Near East. However, further work provided much stronger evidence that more than three-quarters of present-day European mtDNAs most likely stemmed directly from indigenous Mesolithic or Palaeolithic ancestors, and also showed that, with a sufficiently large sample size and a better resolved phylogeny [41], putative clades of mtDNA exhibit gradients similar to those of other marker systems [31,32].

This conclusion is supported by some analyses of the paternally transmitted Y-chromosome [42,43] but not by other studies that inferred admixture coefficients considering different potential parental populations [44]. In the context of this debate, the dissection of one minor autochthonous European mtDNA clade, haplogroup V, and analyses of its geographic distribution [45,46] constituted the first clear evidence that Lateglacial expansions of Palaeolithic populations from refuge areas in Southern Europe could have had a major impact in repopulating the continent. Evidence of the overwhelming importance of the Franco-Cantabrian glacial refuge has been provided by dissecting haplogroups H – the most common haplogroup in Europe (40%–50% of mtDNAs) – and U5 into sub-haplogroups. The age estimates and geographic distributions of H1, H3, V and U5b (Figure 3) strongly support a repopulation of much of Western and Northern Europe from the southwestern refuge zone in the Late Upper Paleolithic [17,20,47,48]. That this Lateglacial expansion must have been dramatic is testified by the dominating star-like structure of

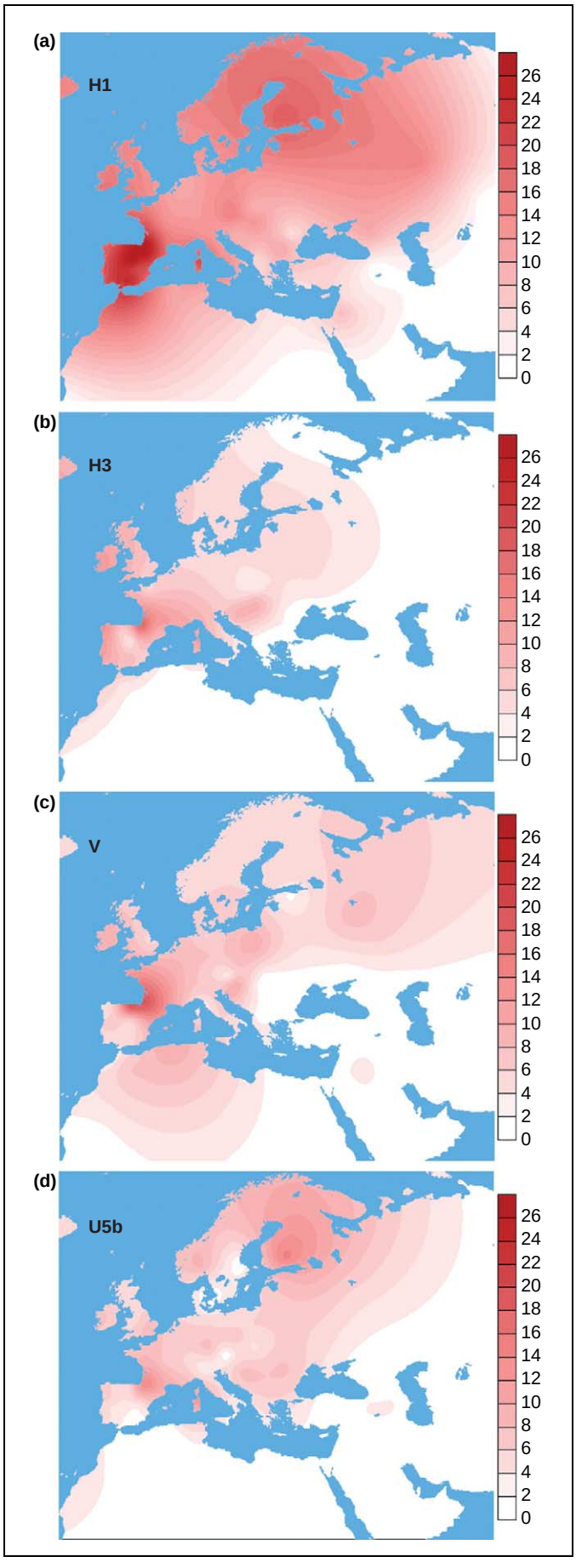

Figure 3. The spatial frequency distributions of haplogroups **(a)** H1; **(b)** H3; **(c)** V; and **(d)** U5b. Data from the Saami population were not included because of their outlier haplogroup frequencies, which are probably caused by profound drift and/or founder events.

sub-haplogroups H1 and H3. It corresponds with an expansion of radiocarbon dates in Western Europe from ~16 000 years ago, and was probably followed by local re-expansions from ~11 500 years ago after the relapse to glacial conditions in the Younger Dryas [49].

On a methodological note, this work has also suggested that a reliable estimation of the different temporal layers of input to a sub-continental mtDNA pool should now take the complete mtDNA phylogeny into full consideration and work upwards from the potentially earliest layers (pioneer settlement) to the most recent layers (post-Neolithic immigration). Having established the basic features of the Palaeolithic mtDNA pool of Europe towards the end of the Pleistocene, one could now turn to the identification of mtDNA lineages that were introduced later to Europe in Neolithic and post-Neolithic times. It seems that all of the major haplogroups (U, JT and R0) now found in Europe were already present in the Late Palaeolithic times, most of which were then represented by their major sub-branches. The small but non-negligible immigration in the Neolithic (and later) seems to have brought some younger sub-haplogroups, such as J1b1, J2a and T1a, and perhaps some that are rare, such as R1, R2 and N1a. The major difficulty in deciding the entry time of all 'founder haplotypes' [41], including the rarer haplotypes, is the enormous amount of sampling that is needed to distil a sufficient number of descendants of the targeted haplotypes across Europe and the Near East. But simplistic admixture analyses that are blind to the different time horizons (not to mention back-migration) and built on unrealistic *ad hoc* hypotheses do not provide a short-cut.

Another way to determine the Neolithic contribution to European mtDNAs has often been envisioned, especially by archaeologists, in the direct analysis of ancient mtDNA from Neolithic remains. The difficulties with ancient DNA of modern humans are well known, and reliable results could only be expected in extraordinary circumstances [50,51]. Recently, Haak *et al.* [52] have apparently triumphed in the 'lottery of ancient DNA studies' (to use the words of Peter Underhill) by detecting variation in a minor haplogroup, N1a, from human remains associated with *Linearbandkeramik* (LBK) pottery across Central Europe that partially matches modern European and Near Eastern haplotypes. This finding, which seems impossible to explain as artefactual, indicates that haplogroup N1a was present in Europe 7000 years ago and suggests that the people involved with the early LBK culture were initially limited to a small area in the middle Danube basin, enabling drift effects to occur, whereas some of their descendants began to migrate and helped the new culture to spread and develop. Extrapolating from these data to the 'Neolithic input' to Europe as a whole would, however, be ill-advised, particularly because the spread of the Neolithic in Europe was not a monolithic and homogeneous process and, of course, 'Neolithic' does not equal 'LBK', temporally or spatially [53].

Therefore, it seems that there is no short-cut for distinguishing between autochthonous and immigrated haplotypes in the modern mtDNA pool at each broad time interval. Only the systematic screening of particular mutations from the complete mtDNA phylogeny in large samples from all over West Eurasia can give us the desired diachronic partition of the mtDNA pool. Results from short stretches of mtDNA drawn from a few samples of old bones and teeth can, at best, only supplement the inferences drawn from present-day complete mitochondrial genomes.

Perspective and concluding remarks

Mitochondrial DNA, along with DNA from the non-recombining portion of the Y chromosome, will be used more and more in elucidating human evolution and pioneer settlement patterns because, at least for the moment, autosomal genes lack the high resolution of the uniparental markers. A prerequisite for future progress is, however, the use of the information from the entire DNA molecule. There is still much work to be done to find all the polymorphic sites on the Y chromosome, but already the SNP set provided by Underhill *et al.* [54,55] has led to a quantum leap in the phylogeographic insights available on the male line of descent. The final level of molecular resolution has already been reached for mtDNA, but the task of determining and interpreting the worldwide variation has just begun, with tens of thousands of mtDNAs to be selected worldwide for complete sequencing in the next few years. Of course, sequencing entire mtDNA genomes is much more expensive and is more labour intensive than the approaches used in the past – but the efforts are paying off, as the resolution of the mtDNA tree, and the phylogeographic resolution that goes with it, has been greatly improved. This can be demonstrated with the poor HVS-I resolution around the European consensus HVS-I sequence, which could ambiguously be allocated to H/HV*/U*/R* [40], but is resolved by complete sequencing [17,20,56]. HVS-I based haplogroups can yield a definitely weaker geographic signal than a comparable number of Y chromosome SNP haplogroups (for more details, see Figure 8.1 of Ref. [57]).

The major obstacle (in addition to technical problems) has always been the reluctance of population geneticists to make full use of complete or partial coding-region data, because this would inevitably entail the adoption of the principles of phylogeography [58] – the systematic comparison of the sequence data under consideration with the current database of complete sequences that comprise the global mtDNA phylogeny, taking into account both geographic and ethnic distributions of haplogroups and sub-haplogroups. However, in medical and forensic genetics, complete mtDNA sequencing is gaining more attention [11,56,59,60]. A new phase of mutual exchange of approaches and insights in which these disciplines would fully interact with population genetics still lies ahead of us [61] – but is, hopefully, just around the corner [62].

Acknowledgements

This research received support from The Royal Society, The British Academy, The Italian Ministry of the University (Fondo Investimenti Ricerca di Base 2001, Progetti Ricerca Interesse Nazionale 2005) and Fondazione Cariplo.

References

1 Brown, W.M. (1980) Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc. Natl. Acad. Sci. U. S. A.* 77, 3605–3609

2 Macaulay, V. *et al*. (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* 64, 232–249

3 Torroni, A. *et al*. (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.* 53, 563–590

4 Torroni, A. *et al*. (1994) Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. *Am. J. Phys. Anthropol.* 93, 189–199

5 Torroni, A. *et al*. (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, 760–776

6 Richards, M. *et al*. (1998) Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* 62, 241–260

7 The Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* 12, 339–348

8 Ozawa, T. *et al.* (1991a) Distinct clustering of point mutations in mitochondrial DNA among patients with mitochondrial encephalomyopathies and with Parkinson's disease. *Biochem. Biophys. Res. Commun.* 176, 938–946

9 Ozawa, T. *et al.* (1991b) Patients with idiopathic cardiomyopathy belong to the same mitochondrial DNA gene family of Parkinson's disease and mitochondrial encephalomyopathy. *Biochem. Biophys. Res. Commun.* 177, 518–525

10 Ozawa, T. *et al*. (1995) Genotype and phenotype of severe mitochondrial cardiomyopathy: a recipient of heart transplantation and the genetic control. *Biochem. Biophys. Res. Commun.* 207, 613–620

11 Tanaka, M. *et al*. (2004) Mitochondrial genome variation in eastern Asia and the peopling of Japan. *Genome Res.* 14, 1832–1850

12 Quintana-Murci, L. *et al*. (1999) Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nat. Genet.* 23, 437–441

13 Kivisild, T. *et al*. (2002) The emerging limbs and twigs of the East Asian mtDNA tree. *Mol. Biol. Evol.* 19, 1737–1751

14 Ingman, M. *et al*. (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* 408, 708–713

15 Maca-Meyer, N. *et al.* (2001) Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet.* DOI:10.1186/1471-2156-2-13 (http://www.biomedcentral.com/1471-2156/2/13)

16 Finnilä, S. *et al*. (2001) Phylogenetic network for European mtDNA. *Am. J. Hum. Genet.* 68, 1475–1484

17 Achilli, A. *et al*. (2004) 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.* 75, 910–918

18 Palanichamy, M.g. *et al*. (2004) Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Am. J. Hum. Genet.* 75, 966–978

19 Friedlaender, J. *et al*. (2005) Expanding Southwest Pacific mitochondrial haplogroups P and Q. *Mol. Biol. Evol.* 22, 1506–1517

20 Achilli, A. *et al*. (2005) Saami and Berbers – an unexpected mitochondrial DNA link. *Am. J. Hum. Genet.* 76, 883–886

21 Kivisild, T. *et al*. (2005) The role of selection in the evolution of human mitochondrial genomes. *Genetics* 172, 373–387

22 Macaulay, V. *et al*. (2005) Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308, 1034–1036

23 Thangaraj, K. *et al*. (2005) Reconstructing the origin of Andaman Islanders. *Science* 308, 996

24 Trejaut, J.A. *et al.* (2005) Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol.* DOI:10.1371/journal.pbio.0030247 (http://biology.plosjournals.org/)

25 Sun, C. *et al*. (2006) The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes. *Mol. Biol. Evol.* 23, 683–690

26 Behar, D.M. *et al*. (2006) The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *Am. J. Hum. Genet.* 78, 487–497

27 Amorim, A. (1999) Archaeogenetics. *J. Iberian Archaeol.* 1, 15–25

28 Renfrew, C. (2000) Archaeogenetics: towards a population prehistory of Europe. In *Archaeogenetics: DNA and the Population Prehistory of Europe* (Renfrew, C. and Boyle, K., eds), pp. 3–11, McDonald Institute Monographs

29 Renfrew, C. (1992) Archaeology, genetics and linguistic diversity. *Man* 27, 445–478

30 Cavalli-Sforza, L.L. (2004) Archaeology, genes and language: reflecting on five decades of human genetics. In *Traces of ancestry: studies in honour of Colin Renfrew* (Jones, M., ed.), pp. 3–10, McDonald Institute Monographs

31 Cavalli-Sforza, L.L. *et al*. (1994) *The History and Geography of Human Genes*, Princeton University Press

32 Richards, M. *et al*. (2002) In search of geographical patterns in European mitochondrial DNA. *Am. J. Hum. Genet.* 71, 1168–1174

33 Richards, M. *et al*. (2003) Analyzing genetic data in a model-based framework: inferences about European prehistory. In *Examining the Farming/Language Dispersal Hypothesis* (Bellwood, P. and Renfrew, C., eds), pp. 459–466, McDonald Institute Monographs

34 Cann, R.L. *et al*. (1987) Mitochondrial DNA and human evolution. *Nature* 325, 31–36

35 Dansgaard, W. *et al*. (1993) Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* 364, 218–220

36 Schulz, H. *et al*. (1998) Correlation between Arabian Sea and Greenland climate oscillations of the past 110 000 years. *Nature* 393, 54–57

37 Kivisild, T. *et al*. (2004) Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. *Am. J. Hum. Genet.* 75, 752–770

38 Kong, Q-P. *et al*. (2003) Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. *Am. J. Hum. Genet.* 73, 671–676

39 Torroni, A. *et al*. (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144, 1835–1850

40 Richards, M. *et al*. (1996) Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 59, 185–203

41 Richards, M. *et al*. (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* 67, 1251–1276

42 Semino, O. *et al*. (2000) The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290, 1155–1159

43 Rootsi, S. *et al*. (2004) Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.* 75, 128–137

44 Dupanloup, I. *et al*. (2004) Estimating the impact of prehistoric admixture on the genome of Europeans. *Mol. Biol. Evol.* 21, 1361–1372

45 Torroni, A. *et al*. (1998) mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* 62, 1137–1152

46 Torroni, A. *et al*. (2001) A signal, from human mtDNA, of postglacial recolonization in Europe. *Am. J. Hum. Genet.* 69, 844–852

47 Loogväli, E-L. *et al*. (2004) Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol. Biol. Evol.* 21, 2012–2021

48 Pereira, L. *et al*. (2005) High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res.* 15, 19–24

49 Gamble, C. *et al*. (2005) The archaeological and genetic foundations of the European population during the Lateglacial: implications for 'agricultural thinking'. *Camb. Archaeol. J.* 15, 193–223

50 Bandelt, H-J. (2005) Mosaics of ancient mitochondrial DNA: positive indicators of non-authenticity. *Eur. J. Hum. Genet.* 13, 1106–1112

51 Gilbert, M.T.P. *et al*. (2005) Assessing ancient DNA studies. *Trends Ecol. Evol.* 20, 541–544

52 Haak, W. *et al*. (2005) Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* 310, 1016–1018

53 Zvelebil, M. (2004) Who were we 6000 years ago? In search of prehistoric identities. In *Traces of ancestry: studies in honour of Colin Renfrew* (Jones, M., ed.), pp. 41–60, McDonald Institute Monographs

54 Underhill, P.A. *et al*. (2000) Y chromosome sequence variation and the history of human populations. *Nat. Genet.* 26, 358–361

55 Underhill, P.A. *et al*. (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.* 65, 43–62
56 Coble, M.D. *et al*. (2004) Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians. *Int. J. Legal Med.* 118, 137–146
57 Forster, P. *et al*. (2004) MtDNA markers for Celtic and Germanic language areas in the British Isles. In *Traces of ancestry: studies in honour of Colin Renfrew* (Jones, M., ed.), pp. 99–111, McDonald Institute Monographs
58 Avise, J.C. (2000) *Phylogeography: the History and Formation of Species*, Harvard University Press
59 Howell, N. *et al*. (2003) Sequence analysis of the mitochondrial genomes from Dutch pedigrees with Leber Hereditary Optic Neuropathy. *Am. J. Hum. Genet.* 72, 1460–1469
60 Carelli, V. *et al*. (2006) Haplogroup effect and recombination of mitochondrial DNA: novel clues from the analysis of Leber hereditary optic neuropathy pedigrees. *Am. J. Hum. Genet.* 78, 564–574
61 Bandelt, H-J. *et al*. (2005) Low 'penetrance' of phylogenetic knowledge in mitochondrial disease studies. *Biochem. Biophys. Res. Commun.* 333, 122–130
62 Bandelt, H-J. *et al*. (2006) *Human Mitochondrial DNA and the Evolution of Homo sapiens*. (*Series* Nucleic Acids and Molecular Biology Vol.18), Springer-Verlag
63 Andrews, R.M. *et al*. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23, 147
64 Watson, E. *et al*. (1997) Mitochondrial footprints of human expansions in Africa. *Am. J. Hum. Genet.* 61, 691–704