

Archaeogenetics — Towards a ‘New Synthesis’?

Guest Editorial

Colin Renfrew

It seems a timely moment to review human population history of the five continents as it emerges from recent archaeogenetic studies, as summarised in the reviews of this special issue of *Current Biology*. Has the ‘new synthesis’ — between genetics, archaeology and linguistics — arrived which I, perhaps incautiously, heralded a few years ago [1]? These highly informative reviews document, it seems to me, both achievement and uncertainty: the achievement relates to the remarkably consistent picture which has now emerged about the out-of-Africa emergence of our own species *Homo sapiens* and the initial peopling of the Earth. The uncertainty involves the application of archaeogenetics to the more recent, Holocene period, when most of the planet was already peopled — except much of Oceania — and sedentary, farming-based communities emerged. Here, it appears that much of our current understanding still depends on archaeological or, sometimes, linguistic evidence. And, with a few exceptions, the archaeogenetic evidence has not yet been assimilated into a genuine synthesis; but, let us begin with the good news.

Out-of-Africa: An Archaeogenetic Success Story

During the nineteenth and much of the twentieth century fossil evidence for early hominins was emerging from many parts of the world. It seemed possible also that the transition from *Homo erectus* (seen in Europe and Asia as well as Africa) to *Homo sapiens* might have involved contributions from the *H. erectus* populations on different continents, or other regional species such as the Neanderthals (*Homo neanderthalensis*) in Europe (the ‘multiregional hypothesis’). A favoured area for the focus of the transition was western Asia. However, since the advent of genetic studies of human evolution, and in particular since the publication of the seminal paper by Cann, Stoneking and Wilson in 1987 [2], the picture, based on mitochondrial DNA (mtDNA) and supported by Y-chromosomal (NRY) DNA, has changed: the transition from *Homo erectus* to *Homo sapiens* took place in Africa, modern humans were established there by some 200 thousand years ago (kya) and the major dispersal (or dispersals) out of Africa — giving rise to essentially all humans living outside Africa today — took place some 70–50 kya.

This fundamental realisation sets a new framework for human history. There is a clear ‘before’ (Africa) and ‘after’ (global) structure to the narrative in the human story. The relatively recent date of our common origin, of the order of 70,000 years, has profound implications. It seems likely that the human capacity for speech is a significant part of the speciation story and was a shared attribute of all sapient humans by 100 kya and perhaps well before. Moreover,

despite modern human adaptation, reviewed here by Pritchard, Pickrell and Coop [3], it seems reasonable to assert that speciation in our species was accomplished some 100 kya, and that most of the behavioural changes seen since then are cultural rather than genetic adaptations. These include the development of farming, of cities, of new technologies, of writing and of a monetary economy — all products of what we may term a tectonic phase in human evolution [4]. Since the onset of this tectonic phase, some 50 kya, the momentum in cultural evolution is no longer driven primarily by genetic differences between populations, although the significance of ‘hard sweeps’ and ‘soft sweeps’ of selection should not be overlooked [3]. Despite the dominant role of culture in driving more recent changes in human behaviour, archaeogenetic research can help us in reconstructing the details of the demographic history of the world. But, of course, also in this area cultural factors are important and can impact genetics: one example, as noted by Mark Stoneking and Frederic Delfin in their review on East Asia [5], is the greater genetic difference between populations for NRY markers than for mtDNA, which reflects a higher female migration rate because of the social practice known as



Figure 1. Genghis Khan.

Genghis Khan left a Y-chromosomal legacy that is still apparent in present-day Asian populations, indicating how social factors such as status can influence the genetic make-up of human populations. (Image: Bridgeman Art Library.)

'patrilocality', whereby women move to the residence of their husbands [6]. Another example is the high male fertility occasionally associated with one specific paternal lineage, for instance the high frequency of an NRY lineage attributed to Genghis Khan and his male descendants [7] (Figure 1). This influence of social and cultural practices upon genetic patterns is perhaps one factor which explains why archaeological interpretations seem to become more difficult after the initial peopling of our planet.

For most continents, except perhaps the Americas, the story of the initial peopling has been triumphantly advanced by progress in archaeogenetics, and has previously been well summarised [8,9]. In south and east Asia, there are indications that the initial 'southern route' dispersal out of Africa may still be recognised in the molecular genetic signature of indigenous Australians and some 'refugia' populations, such as the Malaysian Seman, Filipino Negritos and the Andamanese (Figure 2). Another impressive contribution is the clear indication of the repopulation of northern Europe after the Late Glacial Maximum (LGM), as well reviewed in this issue by Soares *et al.* [10]. It is notable also that the picture in Oceania is becoming clearer, with the 'second occupation', as Manfred Kayser in his review on Oceania [11] terms it, originating in East Asia (possibly Taiwan) and involving a migration to island Melanesia and there developing the Lapita cultural complex and the Proto-Oceanic language before expanding to Polynesia in a further series of migrations.

For the Americas, which are discussed by Dennis O'Rourke and Jennifer Raff [12], however, the jury is still out on the date of the earliest colonisation. There, the underlying difficulty is that the land mass of Beringia through which migration into the Americas took place became depopulated during the LGM, so that the features of its population before that time are difficult to determine. It remains odd, however, that the earliest archaeological dates for settlement in the Americas come from South America, with the site of Monte Verde in Chile well documented around 14,000 BP [13]. The very early dates claimed for the human occupation at Pedra Furada in Brazil [14,15], before 30 kya and long before the LGM, are not accepted by many scholars, but still leave one uneasy that there may be more to learn. This is an area where work is continuing, and perhaps difficulties in obtaining informed consent for samples from some Native American groups have delayed progress. Until there is consensus about the initial colonisation of the Americas and its genetic imprint, it may be difficult to achieve a more effective archaeogenetic narrative for the remainder of the pre-Columbian period — in particular the population histories associated with sedentism and farming in Central America and Peru, as well as the emergence of complex societies in the Americas.

Current Problems

It is indeed with the secondary occupations, often involving the coming of farming and the demographic increases associated with sedentism, that problems arise. This is certainly the case in South Asia, discussed in the review by Partha Majumder [16], where a clear distinction between tribal and caste societies is being claimed, with the tribals being regarded as the 'autochthones of India'. At the same time, it is unclear whether these populations can be declared descendants of the earliest settlers of south Asia, nor is it at all clear precisely when the caste system originated. The caste system may be something which developed over the



Figure 2. Negritos.

Negritos of the Philippines are possibly an example for 'refugial' human populations that may be remnants of the initial spread of humans out of Africa. (Image: Time & Life Pictures/Getty Images.)

past 3000 years, and its origins may be associated with the development of the Hindu religion. But farming-based societies, with which 'tribal' societies are sometimes contrasted, go back very much earlier, and may have been major contributors to the Indian gene pool coming from western and Central Asia. From a linguistic point of view, it is widely supposed that Proto-Indo-European or early Indo-European language, which is ancestral to Vedic Sanskrit and to most of the languages of North India and Pakistan (but not the Dravidian languages of the south), must have come to the sub-continent during the second millennium BC, presumably associated with some incoming population. But, even leaving linguistic issues aside, molecular genetic indicators for these migrations have not been very clearly identified, and it would be useful to know more of origins of these farming populations, which are often termed 'caste societies'.

In Europe, comparable problems arise after the first colonisation of the Upper Palaeolithic (Figure 3). This process, as Soares and colleagues show [10] seems well reflected in

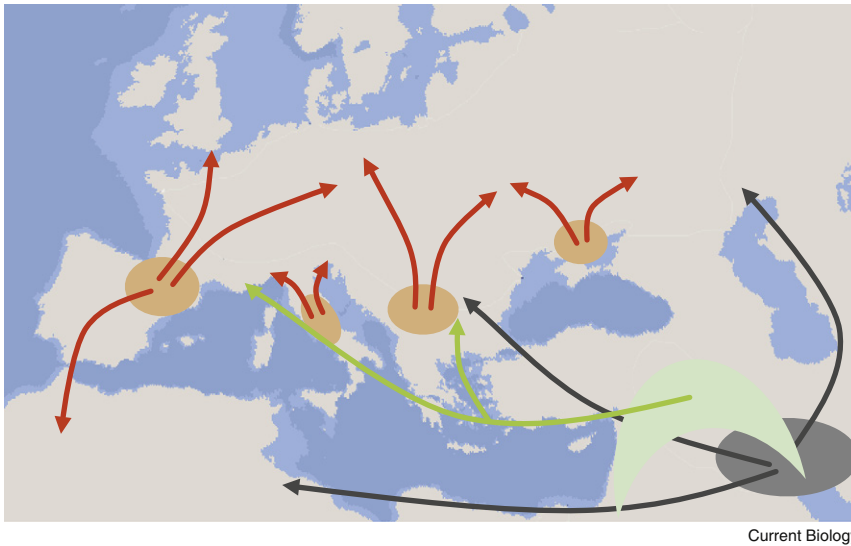


Figure 3. European migrations.

This schematic map depicts major migratory events thought to have affected the gene pool of modern Europeans. Black arrows indicate the first settlement by modern humans around 45 thousand years ago (kya). At the end of the last ice age, around 10–15 kya, Europe was re-populated from glacial refugia (red arrows). Around 8–10 kya, Neolithic farmers came to Europe from Anatolia and the Fertile Crescent (green arrows). (Figure by Alessandro Achilli and Antonio Torroni.)

mtDNA haplogroups U5 and U8, and the haplogroups associated with the post-LGM recolonisation, including H5 and U5b3. It is indeed highly satisfactory that these can be assigned a more precise chronology. But Soares *et al.* [10] are candid in pointing out that this molecular genetic chronology cannot yet be securely correlated with the archaeological evidence for the cultural sequence, which is based on toolkits (Aurignacian, Gravettian, Solutrean, Magdalenian etc.). At later stages, however, the picture becomes significantly more fuzzy. They suggest that fresh analyses of almost 2000 complete European mtDNAs provide a chronology suggesting post-glacial rather than Late Glacial expansion times for most of the lineages spreading northwards from south-west Europe at the end of the Younger Dryas phase, around 11–11.5 kya. As they note “*the Mesolithic and its antecedents appear to have left by far the largest mark on present-day mtDNA and Y-chromosome variation*” [10]. This leaves them with very little to say about the following 10000 years of European history and prehistory, including the Near Eastern Neolithic component. Here, they also refer to the very different conclusions of Chiki and colleagues [17], who, using nuclear DNA markers, would see a largely Neolithic ancestry of the European gene pool. But the subsequent demographic history is not discussed by Soares *et al.* [10]. So, although they have interesting speculations to offer about the linguistic history of Europe, these are not based upon archaeogenetic observations. They go so far as to refer to the 3rd millennium BC as “*a time window little explored by archaeogeneticists*” [10]. Here, then, is a time-range where more information is needed.

In addition, there is one obvious path of investigation that does deserve to be explored further — ancient DNA. Of course this is a difficult field — it depends upon the availability of adequately preserved human remains, and the problems of contamination from living humans are well known. But there are puzzling findings from early farming (*Linienbandkeramik*) sites in central Europe, indicating that the populations in question did not survive or at least did not get fully integrated into succeeding populations [18]. This requires further examination, as it bears on a general problem of archaeogenetics. Possible population extinctions, as suggested by this study [18], might call into question the extent to which mtDNA or NRY DNA data from the

contemporary populations represent the communities existing in the relevant locations at the times in question. This is where ancient DNA may yet prove to be of crucial importance; not in establishing detailed patterns for early populations — the data are unlikely to be rich enough for that — but in offering spot checks on the conclusions about the past which we are deriving from data taken from populations living today. This may indeed be where the future lies if speculations based on the phylogeography of haplotypes from currently living populations are to be rooted in historical reality.

Time Depth Is Crucial

There are many other issues of interest in these fascinating reviews. For instance, the treatment of Africa in the review by Campbell and Tishkoff [19] notes the remarkably long duration of the correlation between genetic and cultural and linguistic similarities in some African populations, such as the African pygmies and several Khoesan-speaking populations, which may have diverged more than 35 kya [20]. Very similar conclusions on this theme have been reached [21,22]. It seems remarkable that groups, which one might refer to as ethnic groups, could have retained linguistic and genetic identities over so long a period of time. Of course, most aspects of those linguistic identities are not available to us as languages change, although the persistence of click consonants in some divergent African hunter-gatherers seems to be one of them that is.

Remarkably, also, these observations are not restricted to the remote periods, equivalent to the periods of initial colonisation and its aftermath in Europe and South Asia. They also extend into the more recent periods. In Africa, genetic signatures of historic and prehistoric migration events are visible in several instances, not least the geographic expansion of the Bantu Niger-Kordofanian speakers from Nigeria and Cameroon first into the rainforests of equatorial Africa and then into eastern and southern Africa within the past 5000 years. So, one is presented here with a number of correlations in Africa between genetic and linguistic prehistory, some of them seemingly going back some 35,000 years. This example perhaps seems to offer more grounds for optimism than some distinguished linguists would allow, also in other continents where questions of time depth in historical linguistics are concerned [23,24]. Indeed, there are indications that the application of phylogenetic methods in this area may prove fruitful [25]. As Manfred Kayser notes in his review of Oceania [11], Bayesian phylogenetic analysis of lexical data undertaken there by Gray, Drummond and

Greenhill [26] revealed a Taiwanese origin of the Austronesian languages about 5200. This can, as he discusses, be related to the molecular genetic evidence, although this is admittedly quite complex. A comparable analysis by Gray and Atkinson [27] applied to the Indo-European languages of Europe and South Asia suggests a time depth of the order of 9000 years for their original divergence, much greater than posited by many Indo-European linguists. It is also much more in harmony with a concomitant spread of farming and language as a mechanism for generating the ultimate geographical extent of this language family [28].

Perhaps the most important general point that can be drawn from the reviews assembled in this special issue might be that we have not yet learnt how to interpret the data very effectively. A number of contributors have commented upon the need for simulation studies, based upon explicit models which might allow the testing of specific scenarios [29], and this is likely to be one of the most important future research directions. Above all, the pace of research is now so fast that new insights are soon likely to become available. These are early days in the field of archaeogenetic research, and I predict that over the next twenty years or so a more coherent synthesis of the data from genetics, archaeology and linguistics is likely to emerge than we can yet envisage.

References

- Renfrew, C. (2000). At the edge of knowability: Towards a prehistory of languages. *Camb. Archaeol. J.* 10, 7–34.
- Cann, R.L.M., Stoneking, M., and Wilson, A.C. (1987). Mitochondrial DNA and human evolution. *Nature* 325, 31–36.
- Pritchard, J.K., Pickrell, J.K., and Coop, G. (2010). The genetic history of human adaptation: hard sweeps, soft sweeps and polygenic adaptation. *Curr. Biol.* 20, R208–R215.
- Renfrew, C. (2007). *Prehistory, the Making of the Human Mind* (London: Weidenfeld & Nicolson), pp. 92–98.
- Stoneking, M., and Delfin, F. (2010). The human genetic history of East Asia: weaving a complex tapestry. *Curr. Biol.* 20, R188–R193.
- Oota, H., Settheetham-Ishida, W., Tiwawech, D., Ishida, T., and Stoneking, M. (2001). Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat. Genet.* 29, 20–21.
- Zerjal, T., Xue, Y., Bertorelle, G., Wells, R.S., Bao, W., Zhu, S., Qamar, R., Ayub, Q., Mohyuddin, A., Fu, S., *et al.* (2003). The genetic legacy of the Mongols. *Am. J. Hum. Genet.* 72, 717–721.
- Forster, P. (2004). Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 359, 255–264.
- Mellars, P. (2006). Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. *Science* 313, 796–800.
- Soares, P., Achilli, A., Semino, O., Davies, W., Macaulay, V., Bandelt, H.-J., Torroni, A., and Richards, M.B. (2010). The archaeogenetics of Europe. *Curr. Biol.* 20, R174–R183.
- Kaysner, M. (2010). The human genetic history of Oceania: near and remote views of dispersal. *Curr. Biol.* 20, R194–R201.
- O'Rourke, D.H., and Raff, J.A. (2010). The human genetic history of the Americas: the final frontier. *Curr. Biol.* 20, R202–R207.
- Dillehay, T. (1997). *Monte Verde: A Late Pleistocene Settlement in Chile, Vol. 2 The Archaeological Context and Interpretation* (Washington D.C.: Smithsonian Institution Press).
- Guidon N. and Arnaud B., The chronology of the New World: two faces of one reality. *World Archaeol.* 23, 167–178.
- Meltzer, D.J., Adovasio, J.M., and Dillehay, T.D. (1995). On a Pleistocene human occupation at Pedra Furada, Brazil. *Antiquity* 68, 695–714.
- Majumder, P.P. (2010). The human genetic history of South Asia. *Curr. Biol.* 20, R184–R187.
- Chikhi, L., Destro-Bisol, G., Bertorelle, G., Pascali, V., and Barbujani, G. (1998). Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proc. Natl. Acad. Sci. USA* 95, 9053–9058.
- Haak, W., Forster, P., Bramanti, B., Matsumura, S., Brandt, G., Tänzer, M., Vilems, R., Renfrew, C., Gronenborn, D., Alt, K.W., *et al.* (2005). Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* 310, 1016–1018.
- Campbell, M.C., and Tishkoff, S.A. (2010). The evolution of human genetic and phenotypic variation in Africa. *Curr. Biol.* 20, R166–R173.
- Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.M., Doumbo, O., *et al.* (2009). The genetic structure and history of Africans and African Americans. *Science* 324, 1035–1044.
- Behar, D.M., Vilems, R., Soodyall, H., Blue-Smith, J., Pereira, L., Metspalu, E., Scozzari, R., Makkan, H., Tzur, S., Comas, D., *et al.* (2008). The dawn of human matrilineal diversity. *Am. J. Hum. Genet.* 82, 1130–1140.
- Quintana-Murci, L., Quach, H., Harmant, C., Luca, F., Massonnet, B., Patin, E., Sica, L., Mougouia-Daouda, P., Comas, D., Tzur, S., *et al.* (2008). Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. *Proc. Natl. Acad. Sci. U S A* 105, 1596–1601.
- McMahon, A., and McMahon, R. (2006). Why linguists don't do dates: evidence from Indo-European and Australian languages. In *Phylogenetic Methods and the Prehistory of Languages*, P. Forster and C. Renfrew, eds. (Cambridge: McDonald Institute for Archaeological Research).
- Renfrew, C. (2000). The problem of time depth. In *Time Depth in Historical Linguistics*, C. Renfrew, A. McMahon, and L. Trask, eds. (Cambridge: McDonald Institute), pp. ix–xiv.
- Renfrew, C., and Forster, P. (2006). Introduction. In *Phylogenetic Methods and the Prehistory of Languages*, P. Forster and C. Renfrew, eds. (Cambridge: McDonald Institute), pp. 1–10.
- Gray, R.D., Drummond, A.J., and Greenhill, S.J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323, 479–483.
- Gray, R.D., and Atkinson, Q.D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426, 435–439.
- Renfrew, C. (2002). The 'Emerging Synthesis': the archaeogenetics of farming/language dispersals and other spread zones. In *Examining the Farming/Language Dispersal Hypothesis*, P. Bellwood and C. Renfrew, eds. (Cambridge: McDonald Institute), pp. 3–16.
- Matsumura, S., and Forster, P. (2008). Introduction. In *Simulations, Genetics and Human Prehistory*, S. Matsumura, P. Forster, and C. Renfrew, eds. (Cambridge: McDonald Institute), pp. 1–6.