

## NEWS AND COMMENTARY

### Population genetics

# DNA from the European Neolithic

G Barbujani and L Chikhi

*Heredity* advance online publication, 24 May 2006; doi:10.1038/sj.hdy.6800852

Where did the Europeans' ancestors come from, and when? Are modern Europeans the direct descendants of the paleolithic hunter-gatherers who settled Europe some 45 000 years ago, or do they mainly derive from the Neolithic farmers who spread from the Levant 10 000 years ago? One would expect archaeologists to discuss these questions, but the last few decades demonstrate that the studies of modern genetic diversity can provide crucial relevant evidence. In a recent issue of *Science*, Haak *et al* (2005) go one step further by typing a stretch of ancient mitochondrial DNA in the largest sample so far from European prehistoric farming communities. The authors successfully amplified and typed 24 mitochondrial sequences from skeletons excavated in 16 Neolithic sites in Germany, Hungary and Austria. Only three Neolithic sequences were previously known, and so this study increases the available data by a factor of 10. However, even beyond the technical achievement, these ancient DNA data could provide important insights in an ongoing controversy.

Past evolutionary processes have left specific signatures in the genes of modern populations, and so genetic data have been used to cast light on the sharp transition found in the archaeological record at the beginning of the Neolithic period. Indeed, artefacts associated with farming technologies first appeared in the Near East 10 000 years ago and then spread North and West until, by 5000 years ago, farming activities were all over Europe. Did this happen because of cultural contacts entailing little migration (cultural diffusion), or because farmers dispersed Westwards, bringing into Europe their technologies along with their genes (demic diffusion)? Clearly, different regions must have experienced different blends of cultural exchanges and migratory movements but it makes sense to ask which alternative is a better representation of the overall process.

Genetic studies have failed to settle the controversy so far, because they have been interpreted in different ways. Allele frequencies (Menozzi *et al*, 1978)

and nuclear DNA polymorphisms (Chikhi *et al*, 1998, 2002; Dupanloup *et al*, 2004) show broad clines across Europe, which fit the expectations of a model of demic diffusion (Ammerman and Cavalli-Sforza, 1984). These clines, paralleling the demographic contribution of Neolithic Near Eastern people estimated from archaeological data (Pinhasi *et al*, 2005), were regarded as evidence of a major dispersal of people in the Neolithic period. However, Richards *et al* (1998, 2000) followed a different approach. They estimated the ages of the main groups of mtDNA haplotypes, or haplogroups, and found that only one of them, haplogroup J, was younger than 10 000 years. They then took the frequency of the J haplogroup, about 20%, as an estimate of the Neolithic contribution to the European gene pool (Richards *et al*, 2000). A rather heated debate followed, and is still continuing.

The predictions of the alternative models, cultural versus demic diffusion, can now be tested against ancient DNA data. To do that, Haak *et al* (2005) performed genetic drift simulations to ask whether they could explain by genetic drift alone the change in frequency of a currently rare (<0.02%) haplogroup (N1a) which was found in as many as six of the 24 (25%) DNA sequences. They concluded that Neolithic farmers originating from the Near East left virtually no descendants in Europe and hence favoured the cultural diffusion model. While one should certainly acknowledge the fact that Haak *et al* used simulations to quantitatively test a hypothesis, the whole study prompts a number of questions, which make the conclusions less convincing than they might first appear.

As noted above, a major part of the controversy has revolved around the age of haplogroups. Since these ages were used by Richards and collaborators to identify 'palaeolithic' and 'neolithic' components of the modern gene pool, one would expect Neolithic specimens to only yield 'neolithic', haplogroup J, sequences. The population genetics prediction, however, is that 'Neolithic' people should have both types of haplogroups. Interestingly,

Haak *et al* (2005) found only one sequence belonging to haplogroup J, with six sequences belonging to the currently very rare N1a haplogroup, and 17 to haplogroups that were termed 'Paleolithic' (Richards *et al*, 2000), such as H, V and K. Surprisingly, this information was not used by the authors, even though it demonstrates that ages of molecules cannot be equated with ages of populations, a point made some time ago by supporters of the demic diffusion model (Barbujani *et al*, 1998). Population genetics theory teaches us that migrating people carry alleles and haplogroups in their genome originating from mutations that occurred before, sometimes long before, the migratory movement started, and inferring from the former the date of the latter is never straightforward. It might be legitimate (although, we think, misleading) to term haplotypes derived from mutations <10 000 years old as 'Neolithic', but the frequency of those haplotypes has little to do with the Neolithic contribution to the European gene pool.

Other issues would need to be further explored. Firstly, haplogroup N1a was found to be common in the Neolithic sample, but is virtually absent from modern European populations. Does this tell us anything on the impact of genetic drift in the first farming communities? Secondly, clines have a direction, but not a date, and both the first Paleolithic settlers and the later Neolithic farmers came to Europe from the Southeast. Therefore, although clines radiating from the Near East have now been shown for hundreds of DNA loci (Belle *et al*, 2006), the possibility exists that these clines were generated (i) by selection (but that should then be limited to some loci); (ii) through a series of founder events during the first colonization of the continent; (iii) during other, so far undefined, processes occurring along a Southeast–Northwest axis, provided these processes affected much of Europe. Another possibility, recently suggested by Currat and Excoffier (2005) is that ascertainment bias in the choice of polymorphic loci could also generate clines. Although there is no evidence of an ascertainment bias for the many STR loci showing gradients in Europe (Chikhi *et al*, 1998; Belle *et al*, 2006), Currat and Excoffier's (2005) very exciting study demonstrates that it is now becoming possible to address complex questions, and infer by simulation the most probable scenarios for the European prehistorical migrations using genetic data.

For the future, we believe that such simulation studies represent the most promising approach. One advantage of such approaches is that they force us to explicitly formulate what were often previously implicit assumptions. They also allow researchers to test the effect of factors that are difficult to quantify exactly, such as prehistoric effective population sizes, rates of gene flow, and mutation rates, on the likelihood of different scenarios. In the near future, they will also allow users to integrate information from the growing body of ancient genetic diversity, in addition to the broad modern data sets (Anderson *et al*, 2005). The application of the full panoply of modern statistical tools to the impressive data set of Haak *et al* promises to disclose new, and previously out-of-reach, aspects of our evolutionary past.

G Barbujani is at the Dipartimento di Biologia, Università di Ferrara, via Borsari 46, I-44100 Ferrara, Italy and L Chikhi is at the UMR CNRS 5174 Evolution et Diversité Biologique, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse cedex 4, France.

e-mail: g.barbujani@unife.it

Barbujani's lab URL: <http://web.unife.it/progetti/genetica/>

- Ammerman AJ, Cavalli-Sforza LL (1984). *The Neolithic Transition and the Genetics of Populations in Europe*. Princeton University Press: Princeton.
- Anderson CNK, Ramakrishnan U, Chan YL, Hadly EA (2005). *Bioinformatics* **21**: 1733–1734.
- Barbujani G, Bertorelle G, Chikhi L (1998). *Am J Hum Genet* **62**: 488–492.
- Belle EMS, Landry PA, Barbujani G (2006). *Proc R Soc B*, in press.
- Chikhi L, Destro-Bisol G, Bertorelle G, Pascali V, Barbujani G (1998). *Proc Natl Acad Sci USA* **95**: 9053–9058.
- Chikhi L, Nichols RA, Barbujani G, Beaumont MA (2002). *Proc Natl Acad Sci USA* **99**: 11008–11013.
- Curat M, Excoffier L (2005). *Proc R Soc B* **272**: 679–688.
- Dupanloup I, Bertorelle G, Chikhi L, Barbujani G (2004). *Mol Biol Evol* **21**: 1361–1372.
- Haak W *et al* (2005). *Science* **310**: 1016–1018.
- Menozzi P, Piazza A, Cavalli-Sforza LL (1978). *Science* **201**: 786–792.
- Pinhasi R, Fort J, Ammerman AJ (2005). *PLoS Biol* **3**: e410.
- Richards M *et al* (1998). *Am J Hum Genet* **59**: 185–203.
- Richards M *et al* (2000). *Am J Hum Genet* **67**: 1251–1276.
- Further Reading**
- Balaresque P, Manni F, Dugoujon JM, Crousau-Roy B, Heyer E (2006). Estimating sex-specific processes in human populations: Are XY-homologous markers an effective tool? *Heredity* **96**: 214–221.
- Barbujani G, Bertorelle G (2001). Genetics and the population history of Europe. *Proc Natl Acad Sci USA* **98**: 22–25.
- Barbujani G, Goldstein DB (2004). Africans and Asians abroad: genetic diversity in Europe. *Annu Rev Genomics Hum Genet* **5**: 119–150.
- Cavalli-Sforza LL, Menozzi P, Piazza A (1993). Demic expansions and human evolution. *Science* **259**: 639–646.
- Excoffier L, Schneider S (1999). Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. *Proc Natl Acad Sci USA* **96**: 10597–10602.
- Goldstein DB, Chikhi L (2002). Human migration and population structure: what we know and why it matters. *Annu Rev Genom Hum Genet* **3**: 129–152.
- Goodacre S, Helgason A, Nicholson J, Southam L, Ferguson L, Hickey E *et al* (2005). Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* **95**: 129–135.
- Hedrick P, Waits L (2005). Conservation genetics: What ancient DNA tells us. *Heredity* **94**: 463–464.
- Morrison DA, Höglund J (2005). Testing the hypothesis of recent population expansions in nematode parasites of human-associated hosts. *Heredity* **94**: 426–434.
- Pinhasi R, Foley RA, Mirazòn Lahr M (2000). Spatial and temporal patterns in the Mesolithic–Neolithic archaeological record of Europe. In: Renfrew C, Boyle H (eds) *Archaeogenetics: DNA and the population prehistory of Europe*. McDonald Institute for Archaeological Research: Cambridge, pp 45–56.
- Richards M, Macaulay V, Torroni A, Bandelt HJ (2002). In search of geographical patterns in European mitochondrial DNA. *Am J Hum Genet* **71**: 1168–1174.
- Sokal RR, Oden NL, Wilson C (1991). Genetics evidence for the spread of agriculture in Europe by demic diffusion. *Nature* **351**: 143–145.