

Strength in numbers

Austin L. Hughes

Every species has a history. Although no chronicler was there to record them or bard to immortalize them, fruitflies, flycatchers and kangaroos have had their epic migrations, their great plagues, their eras of prosperity and population growth, and their periods of decline.

For our own species, we give the name 'prehistory' to all of our history that occurred before written records. In this sense, most of the history of other species is also prehistoric, and until quite recently we had little means of knowing about it. Only a few of our fellow life-forms — such as domestic animals and our major pests — have found their way into our written records or left traces at our archaeological sites. Most non-human species have lived their histories away from us and have left no artefacts for us to find out about them. Our prospects of reconstructing much of our own prehistory used to be little better. Archaeological evidence is often ambiguous — for example, when technologies or styles of art change, was this because new techniques had been learned or because another culture had replaced the previous one? Even the written record poses as many historical mysteries as it resolves, as we cannot distinguish

between propaganda and objective accounts.

Molecular-biology data offer the promise of at last unlocking the prehistories of our own and other species. In the case of ourselves and a few other organisms, we now have an invaluable resource — a complete genome sequence. Using this sequence as a guide, we can re-sequence from different populations a substantial number of loci sampled from throughout the genome.

Should we believe it when we are told that genetic markers can be used to reconstruct all sorts of ancient events, from the origin of modern *Homo sapiens*, to the spread of agriculture, to the peopling of oceanic islands? It is true that genetic markers — unlike ancient chroniclers — do not lie. But their interpretation raises many a thorny problem and can be as perilous as attempts to decipher ancient inscriptions in an unknown tongue.

Consider the most straightforward type of question that we might want to answer about a species' history. Suppose we want to date the separation of two populations of the same species, divided by a geographical barrier that prevents or substantially reduces gene flow between them. The separation might correspond to a major migration — for example, the first migration of modern humans out of Africa. Naively, we might assume that answering this question would be very simple. We could sequence alternative forms of a gene (alleles) from an individual in each of the two populations and count the base-pair differences between the two sequences. Given an estimate of the mutation rate and the assumption that this rate is roughly constant over time, it is straightforward to estimate the age of the common ancestor of the two sequences.

But this date does not necessarily correspond to the most recent common ancestor of the two populations, because the ancestral population itself would have contained genetic differences. Just by chance, an allele that existed in the ancestral population might have ended up being fixed in one of the descendant populations, while another allele was fixed in the other descendant population. The common ancestor of these two alleles will be significantly more ancient than the time at which the two populations separated, especially when balancing selection acts to maintain such a polymorphism.

In vertebrates, the best-studied case of an ancient, balanced polymorphism is the highly polymorphic loci of the major histocompatibility complex (MHC) of the immune system, in humans generally termed HLA (for human leukocyte antigen). Polymorphic alleles at HLA loci — or, more accurately, allelic lineages — can be very ancient, even pre-dating the most recent common ancestor of humans

Genetic markers

Genetic markers — unlike ancient chroniclers — do not lie. But their interpretation can be as perilous as attempts to decipher ancient inscriptions in an unknown tongue.

and chimpanzees. It now seems that balancing selection at other loci in the human genome may be more common than previously thought — there is probably a balanced polymorphism at the dopamine-receptor D4 locus that could relate to behavioural differences between people with different alleles for this gene. The melanocortin 1 receptor locus, which is involved in pigmentation of skin and hair, is another surprisingly polymorphic site at which balancing selection may operate.

The obvious solution is to sample as many loci as possible. (It is worth noting in this context that the entire mitochondrial genome — the workhorse of genetic marker studies to date — represents only a single locus, as it is inherited as a unit without recombination.) If gene flow between two populations has been completely eliminated, the set of loci that shows the most recent ancestor for the two populations must correspond to loci that were monomorphic at the time of their separation. However, in many cases, gene flow between isolated populations is reduced but not eliminated. Population migrations may follow episodic patterns, as in the multiple migrations out of Africa over hundreds of thousands of years that have recently been proposed for our own species. Given such a complex history, there is a real danger that polymorphism pre-dating population subdivision will be taken as evidence of a still more ancient migration.

Data on worldwide patterns of sequence polymorphism at hundreds or even thousands of loci provide the statistical power to discriminate between chance patterns observed at one or a few loci, and can hence reveal true population histories. In the world of genetic markers, there is definitely strength in numbers. ■

Austin L. Hughes is professor of biological sciences and director of the Biotechnology Institute, University of South Carolina, Columbia, South Carolina 29208, USA.

FURTHER READING

- Ding, Y. C. *et al. Proc. Natl Acad. Sci. USA* **99**, 309–314 (2002).
 Harpending, H. C. *et al. Proc. Natl Acad. Sci. USA* **95**, 1961–1967 (1998).
 Takahata, N. *et al. Mol. Biol. Evol.* **18**, 172–183 (2001).
 Nei, M. *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York, 1987).
 Templeton, A. R. *Nature* **416**, 45–51 (2002).



Tell-tale signs: like this Bronze Age tablet, genetic polymorphisms can offer clues to human history.