Heredity (2010) 104, 419–420

www.nature.com/hdy

NEWS AND COMMENTARY

Mitochondrial phylogeography

The worm in the fruit of the mitochondrial DNA tree

F Balloux

Heredity (2010) **104,** 419–420; doi:10.1038/hdy.2009.122; published online 16 September 2009

Let us assume I gave a seminar. I would tell the audience about my latest results on the population history of the pigmy shrew. My findings would be based on a stretch of DNA comprising several metabolic genes, showing no signs of genetic recombination. Armed with sequences from a large number of individuals sampled over a broad geographical area, I would make some inference on the colonization routes and times. To make life easier, I would restrict my analysis to the mutations I liked best, with nice names having been given to related sequences, rather than relying on dull mathematical quantities. As I reach one of the key conclusions of the lecture, which would go as follows: *'It is obvious from the distribution of haplotypes Amanda, Eugenie* and Hector_2α that the Outer Hebrides were colonised about 50,000 years ago, this was followed by considerable population fluctuations, a bottleneck during the last Ice Age, a swift recovery and a dramatic recent expansion over the last 200 years and…'*. Imagine that, at that climactic stage I was interrupted by someone in the audience. The impertinent would say, '*Sir, can I just ask you whether this confidence in your conclusions may not be misplaced; your analysis is based on a single genetic marker, which comprises genes with a central role in metabolism and is thus likely to have been affected by natural selection*'. An awkward silence may ensue, as I would find it difficult to dismiss this criticism easily.

There are parallels between this hypothetical story and a large body of work using mitochondrial (mt)DNA polymorphisms to reconstruct the past history of innumerable species, including our own. Yet, no one in the audience seems keen to interrupt the speaker. I can only conclude that the limitations of mtDNA as a genetic marker are not fully recognized. This is arguably not because no one has previously expressed such misgivings. Excessive reliance on uniparental markers (that is, mtDNA and Y chromosome) has been criticized before (Ballard and Whitlock, 2004; Pakendorf and Stoneking, 2005; Bazin *et al*., 2006). Arguably, the most scathing assessment was reached by Harpending (2006) who described the entire field as a series of anecdotes ranging from the plausible and interesting to the absurd. It does not help that human phylogeography has become largely divorced from the rest of population genetics and relies on some particularly arcane jargon. As none of these earlier criticisms seem to have had much effect in stemming the tidal wave of papers using mtDNA to make fine demographic inference, I feel compelled to summarize once again what the problems are and why they cannot be swept under the carpet indefinitely.

Unlinked genetic markers trickle down through species' pedigrees independently from one another. Owing to the high stochasticity of population demography, some genetic markers will be reflective of the populations' history and some would not. mtDNA and the Y chromosome do not recombine and therefore represent only a single realization of the many possible outcomes within a given demographic history, irrespective of the number of polymorphic sites typed (Ballard and Whitlock, 2004). As a consequence, the phylogenetic tree of uniparental markers may or may not be informative on the demography of the populations studied. Although it has been said forcefully before that gene trees should not be equated to species or population trees (Nichols, 2001), it seems that this subtle yet important distinction is rarely made. A frankly baffling trend from a population genetics perspective is the apparent increase of papers considering only a single haplogroup, as the problems with demographic stochasticity will be exacerbated even further. This demographic variance is not a major issue when one has access to a reasonable number of autosomal markers, which can be treated as replicates of the same process and averaged over loci to make inference on the population history.

A crucial assumption is that the genetic markers used to make inference on the past history of populations are evolving neutrally; that a non-recombining stretch of DNA comprising 37 genes should be neutral seems a bold hypothesis. Detecting natural selection in non-recombining DNA is difficult. Despite this, there is evidence for natural selection on mtDNA in various taxa (for example, Ballard *et al*., 2007; Fontanillas *et al*., 2005; Oliveira *et al*., 2008), with temperature being often invoked as the likely selective force. The situation in humans is still far from clear. There have been claims based on ratios of synonymous versus non-synonymous mutations (dN/dS ratios), and to a lesser extent, the evolutionary persistence of mutations that human mtDNA may have been affected by climate (Torroni *et al*., 2001; Mishmar *et al*., 2003; Ruiz-Pesini *et al*., 2004). This has been refuted by other studies, which concluded that human mtDNA sequence variation has not been significantly shaped by natural selection (Elson *et al*., 2004; Kivisild *et al*., 2006; Amo and Brand, 2007; Ingman and Gyllensten, 2007). However, all these results (both for and against selection) are questionable as dN/dS ratios are generally inadequate tests for natural selection when working over limited evolutionary time scales at a within-population level (Rocha *et al*., 2006; Kryazhimskiy and Plotkin, 2008). Interestingly, the most comprehensive of the studies concluding that human mtDNA was evolving neutrally did actually highlight the single-nucleotide polymorphism at position 10 398 as a target for natural selection (Kivisild *et al*., 2006). This single-nucleotide polymorphism was also identified again in study, which showed in addition that mitochondrial diversity correlated with the temperature to which human populations were exposed (Balloux *et al*., 2009). Finally, the same single-nucleotide polymorphism has been shown to affect mitochondrial matrix pH and mitochondrial calcium dynamics (Kazuno *et al*., 2006).

There is even a sensible biological explanation why temperature may affect mtDNA sequence variation (Coskun *et al*., 2003; Mishmar *et al*., 2003). As shivering frantically to keep warm is not such a pleasant prospect, we have to rely largely on the heat generated by the oxidative phosphorylation (OXPHOS) cycle, which comprises 13 genes encoded by the mitochondrial genome. The primary

function of the OXPHOS cycle is the synthesis of ATP, the molecular currency of all cells. In a tropical climate where keeping the body warm is not a major issue, the OXPHOS cycle should be optimized toward the production of ATP molecules. As with any other physical process, higher efficiency means less heat is produced; therefore, in colder climates it may make evolutionary sense to trade some energy for heat. The extent to which the distribution of human mtDNA variation has been shaped by environmental factors remains to be assessed. However, given the pivotal role of the OXPHOS cycle in basal metabolism, selective neutrality of mtDNA variation seems a largely untenable *a priori* assumption at least in humans, and probably for any other endothermic species exposed to a wide array of temperatures. Many autosomal markers are undoubtedly affected by selection too. However, this is not expected to be a major issue when multiple unlinked markers are considered, as the signature of selection will be averaged out.

Despite mitochondrial sequence variation covarying with climate in humans (Balloux *et al*., 2009), there are better ways to measure temperature. And, I would argue there are also better genetic markers than mtDNA to infer past population history. I fully appreciate that mtDNA has given us some of the most fundamental results on human evolution at a time when using mtDNA was the only realistic option at hand. I do not question the value of mtDNA in forensics and pedigree reconstruction. It is also likely to remain a valuable tool for inference at a localized geographical scale, particularly when testing specific hypotheses rather than making quantitative inferences on the age or size of the populations studied. It is convenient to type and analyse, and its use in humans raises no serious ethical or societal issue. But all these qualities do not counterbalance the fact that a single locus likely to be under selection is inappropriate for population inference at large geographical scales (or over long periods of time in the context of ancient DNA analysis). We have reached an era in which publicly available data sets of large numbers of complete human genomes are a tangible prospect, and I believe it is now time to move on from the excessive reliance on uniparental markers. Exploiting these new resources of autosomal variation will present significant challenges, but it will not help overcoming them if a large fraction of the community of human population biologists persists in sticking to mtDNA as the marker of choice.

Acknowledgements

I am grateful to Mark Achtman and Richard Nichols for inspiring discussions and encouragements to go ahead with this comment. I also acknowledge the Biotechnology and Biological Sciences Research and the Medical Research Council for financial support.

Dr F Balloux is at the Department of Infectious Disease Epidemiology, MRC Centre for Outbreak Analysis and Modelling, Imperial College Faculty of Medicine, St Mary's Campus, Norfolk Place, London W2 1PG, UK.

e-mail: fballoux@imperial.ac.uk

Amo T, Brand MD (2007). Were inefficient mitochondrial haplogroups selected during migrations of modern humans? A test using modular kinetic analysis of coupling in mitochondria from cybrid cell lines. *Biochem J* **404**: 345–351.

Ballard JWO, Melvin RG, Katewa SD, Maas K (2007). Mitochondrial DNA variation is associated with measurable differences in life-history traits and mitochondrial metabolism in Drosophila simulans. *Evolution* **61**: 1735–1747.

Ballard JWO, Whitlock MC (2004). The incomplete natural history of mitochondria. *Mol Ecol* **13**: 729–744.

Balloux Fo, Handley L-JL, Jombart T, Liu H, Manica A (2009). Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proc R Soc B Biol Sci* **276**: 3447–3455.

Bazin E, Glemin S, Galtier N (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**: 570–572.

Coskun PE, Ruiz-Pesini E, Wallace DC (2003). Control region mtDNA variants: longevity, climatic adaptation, and a forensic conundrum. *Proc Natl Acad Sci USA* **100**: 2174–2176.

Elson JL, Turnbull DM, Howell N (2004). Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am J Hum Genet* **74**: 229–238.

Fontanillas P, Depraz A, Giorgi MS, Perrin N (2005). Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, Crocidura russula. *Mol Ecol* **14**: 661–670.

Harpending H (2006). Anthropological genetics: present and future. In: Crawford MH (ed). *Anthropological Genetics Theory, Methods and Applications*. Cambridge University Press: Cambridge.

Ingman M, Gyllensten U (2007). Rate variation between mitochondrial domains and adaptive evolution in humans. *Hum Mol Genet* **16**: 2281–2287.

Kazuno A-a, Munakata K, Nagai T, Shimozono S, Tanaka M, Yoneda M *et al*. (2006). Identification of mitochondrial DNA polymorphisms that alter mitochondrial matrix pH and intracellular calcium dynamics. *PLoS Genet* **2**: e128.

Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K *et al*. (2006). The role of selection in the evolution of human mitochondrial genomes. *Genetics* **172**: 373–387.

Kryazhimskiy S, Plotkin JB (2008). The Population Genetics of dN/dS. *PLoS Genet* **4**: e1000304.

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S *et al*. (2003). Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA* **100**: 171–176.

Nichols R (2001). Gene trees and species trees are not the same. *Trends Ecol Evol* **16**: 358–364.

Oliveira DCSG, Raychoudhury R, Lavrov DV, Werren JH (2008). Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp nasonia (hymenoptera: pteromalidae). *Mol Biol Evol* **25**: 2167–2180.

Pakendorf B, Stoneking M (2005). Mitochondrial DNA and human evolution. *Annu Rev Genomics Hum Genet* **6**: 165–183.

Rocha EPC, Smith JM, Hurst LD, Holden MTG, Cooper JE, Smith NH *et al*. (2006). Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J Theor Biol* **239**: 226–235.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* **303**: 223–226.

Torroni A, Rengo C, Guida V, Cruciani F, Sellitto D, Coppa A *et al*. (2001). Do the four clades of the mtDNA haplogroup L2 evolve at different rates. *Am J Hum Genet* **69**: 1348–1356.