5 Klein, N.K. and Payne, R.B. (1998) Evolutionary associations of brood parasitic finches (*Vidua*) and their host species: analyses of mitochondrial restriction sites. *Evolution* 52, 299–315
6 Lade, B.I. and Thorpe, W.H. (1964) Dove songs as innately coded patterns of specific behaviour. *Nature* 202, 366–368
7 Beecher, M.D. (1996) Birdsong learning in the laboratory and field. In *Ecology and Evolution of Acoustic Communication* (Kroodsma, D.E. and Miller, E.H., eds), pp. 61–78, Cornell University Press
8 ten Cate, C. and Vos, D.R. (1999) Sexual imprinting and evolutionary processes in birds: a reassessment. *Adv. Study Behav.* 28, 1–31
9 Grant, P.R. and Grant, B.R. (1997) Hybridization, sexual imprinting, and mate choice. *Am. Nat.* 149, 1–28
10 Irwin, D.E. and Price, T. (1999) Sexual imprinting, learning and speciation. *Heredity* 82, 347–354
11 Owens, I.P.F. *et al.* (1999) Sexual selection, speciation and imprinting: separating the sheep from the goats. *Trends Ecol. Evol.* 14, 131–132
12 Laland, K.N. (1994) On the evolutionary consequences of sexual imprinting. *Evolution* 48, 477–489
13 ten Cate, C. and Bateson, P. (1988) Sexual selection: the evolution of conspicuous characteristics in birds by means of imprinting. *Evolution* 42, 1355–1358

Human mitochondrial DNA recombination: can it be true?

Pity the evolutionary geneticists trying to keep track of mitochondrial DNA (mtDNA) inheritance. Sure there are rules of thumb, but none has ever been so sturdy or so widely applicable as Mendel's Laws are for eukaryote nuclear genes. Plants exhibit the panoply of mtDNA inheritance, with some showing strictly sex-limited transmission (either paternal or maternal) and a few showing biparental inheritance[1]. For other anisogamous organisms, the working model has usually been one of strictly maternal inheritance, although now that we have well confirmed, striking exceptions in mussels[2,3], as well as occasional reports of paternal leakage in other organisms[4–6], one might never feel too sure. The most recent affront to simplicity is the claim that a human's mtDNA is sometimes a mixture of conventional maternal mtDNA that has recombined with mtDNA from somewhere else[7–9]. The claims are disconcerting because so many researchers study human mtDNA, and do so while relying upon the assumption that non-maternal inheritance and recombination do not occur.

The first reports suggesting recombination focused on the apparent implausibility of mutation as the sole cause of homoplasy[7,8]. Human mitochondria are famously homoplastic, particularly in the control region of the circularly mtDNA genome[10–12]. It has long been assumed that our inability to generate a single evolutionary gene tree, which explains each polymorphism via a single mutation, was a simple consequence of many polymorphisms having been caused by multiple mutations at the same base position. Hagelberg *et al.*[7] found a polymorphism that appeared to have arisen more than once among pacific islanders, but had never been reported from humans in other geographic regions. However, that observation was recently found to be in error and must be discounted[13]. Eyre-Walker *et al.*[8] avoided control region data, and focused instead on synonymous polymorphisms (i.e. in protein-coding regions but not affecting amino-acid sequences) that were found around the entire mtDNA circle. They inquired just how much of the considerable homoplasy, found for these synonymous polymorphisms, could be accounted for by mutation models. Their conclusion was arrived at through the backdoor – in finding that mutation models were insufficient to explain homoplasy, the authors were left with some kind of recombination as the only remaining explanation.

Of course, this sort of backdoor conclusion is fine when all that comes in is an otherwise reasonable notion. But, recombination in human mitochondria is not a reasonable notion – at least in so far as it flies in the face of what other knowledge we have of mtDNA inheritance in humans, and in so far as human evolutionary geneticists have assiduously assumed that it cannot happen. When I first heard Eyre-Walker describe his and colleagues' findings, I was impressed by the clear logic behind their assessment of the mutational component of homoplasy, but I simply was not prepared to accept a conclusion of recombination. I figured that we are fairly ignorant of mutation, and supposed that it might vary over time and base positions in ways that are too awkward for any analysis with a small number of parameters. As complicated as such a process might be, it seemed to be at least as parsimonious as an invocation of recombination.

But however reasonable such complacency, it does not easily weather a more recent report by Awadalla, Eyre-Walker and Maynard Smith[9]. Awadalla *et al.* examined linkage disequilibrium (LD) between pairs of polymorphic sites, as a function of the distance between polymorphic sites. Linkage disequilibrium generally arises as a simple by-product of mutation, linkage and genetic drift, and it can be increased or decreased depending on these and other evolutionary factors. Of course it will decrease, on average, the more recombination occurs between the pairs of polymorphic sites upon which it is measured. This simple idea, that recombination will generate a negative correlation between LD and DNA distance, has been used several times to assess the historical role of recombination, particularly in contexts where conventional assessments of recombination are difficult[14–17]. Unlike most tests of recombination[18,19] that are at risk of misconstruing recurrent mutation as recombination, a pattern of decreasing LD is not expected to mislead in this way. To help see this, consider the basic nonparametric protocol employed by Awadalla *et al.* First, each of all the possible pairs of informative polymorphisms (i.e. those in which the rarer base occurs at least twice in the sample of sequences) are subjected to two calculations: one that gives a measure of LD and another that simply counts the number of base pairs in between the two positions. Second, the correlation is measured between these paired variables. Third, the positions of the same polymorphisms are randomly scrambled along the sequence and the correlation is reassessed for the randomized data. Finally, this scrambling is repeated many times to generate a distribution of correlation coefficients with which to assess the significance of the actual correlation. Given this design, it is hard to see how the researchers could be misled by mutation, no matter how recurrent or changing one might suppose mutation to be. So long as different mutations occur independently of one another, it is hard to see how they could generate a negative correlation between polymorphism and distance.

Given that a significantly negative correlation was found in several independent data sets, from humans and chimpanzees[9], it seems unlikely that the reported correlations are a fluke or are

 0169-5347/00/$ – see front matter PII: S0169-5347(00)01832-2

caused by an erroneous data set. But, that does not mean that a skeptic does not have room to work. Rather, an investigator who suspects a mistaken conclusion must find some other mechanism, besides recombination, that could cause covariation in polymorphism patterns as a function of the distance between them (specifically a negative covariation). Here are two general possibilities. First, mutation might indeed occur in nonindependent ways. For example, mutations might sometimes occur in pairs or multiples and, when they do, they may be more likely to occur near one another, rather than further apart. There is a wealth of evidence against models of this sort, going back at least to the classic Luria–Delbruck experiments[20], but they do not rule out occasional bursts of mutation in mtDNA. Second is that data recording might introduce nonindependent errors. For example, suppose that those who did the DNA sequencing tended to gather data in parallel (i.e. all samples for one segment of the mtDNA and then all samples for the next), rather than in series (one complete and then the second complete, etc.). Parallel sequencing protocols are the rule, and it is possible that when errors do occur they are introduced in such a way as to cause some covariation with distance. The problem with this notion is that the results of Awadalla *et al.* rely upon informative polymorphisms, which turn up multiple times in the sample. By contrast, most DNA sequencing errors are likely to be singletons (i.e. they occur only once in a data set).

If Awadalla *et al.* are correct, then the actual process of exchange among mtDNAs is unlikely to be anything resembling the crossing over that goes on during meiosis in the nucleus. The authors made two suggestions for what might actually be occurring. First is the possibility that in the zygote or early embryo, some mtDNA that comes in with the sperm gets taken up by maternal mitochondria and somehow replaces a homologous segment in the maternal mtDNA. Ordinarily, the paternal mitochondria and mtDNA are degraded[21], but perhaps some fragments of DNA survive to be incorporated. A second possibility is that the exchange is actually between copies of the mtDNA (or portions thereof) that moved into the nucleus at some earlier date, and then somehow moved back and became incorporated in the mtDNA. As unlikely as both of these explanations might seem, it might be that an unlikely explanation is just what is called for. Although Awadalla *et al.* do not provide estimates of the rates or sizes of gene transfer tracts, it might be that a few rare events could explain the pattern that has been described.

So why should we care if there is some recombination among mtDNAs? For evolutionary biologists, the concern is that a true history that includes recombination would obviate the ubiquitous assumption of a bifurcating gene tree history. Any DNA, mitochondrial or otherwise, that does not engage in recombination will necessarily have a branching history (although it might only be revealed to us by mutations that distinguish sequences and permit inferences of history). But if recombination has occurred then the true history is literally a complex network and not a branching tree at all. In fact, human mtDNA researchers can rarely generate a particular gene tree estimate of any confidence – as mentioned above, homoplasy is pervasive. Nevertheless, we assume that a gene tree is the correct model and thus discussions are often focused about the timing of particular parts on a gene tree, such as the most recent common ancestor of all the items in a sample. Such talk does not make sense for a sample of DNAs with a recombinogenic history. Some other particular analyses would also be undermined, such as those that rely upon the distribution of pairwise differences, which assume a single branching history[22].

Fortunately, even if recombination is in our mtDNA history, all would not be lost. One of the most awkward features of mtDNA history, and one that is typically overlooked, is that it is but one of many thousands of genealogical histories that our gene pool has experienced. Different genes have different histories and the variance among genes can be enormous, just by chance. Thus, the mtDNA was always bound to be atypical, just as would any gene, even without the action of something deterministic, such as natural selection. If recombination has been occurring, then the mtDNA has a more complicated history, but one that can be expected to be closer to that for the average of all genes. The history might be harder to decipher, but what is found might be a more accurate estimate of ancient human history.

Jody Hey

Dept of Genetics, Rutgers University, Nelson Biological Labs, 604 Allison Road, Piscataway, NJ 08854-8082, USA (jhey@mbcl.rutgers.edu)

References

1 Mogensen, H.L. (1996) The hows and whys of cytoplasmic inheritance in seed plants. *Am. J. Bot.* 83, 383–404

2 Hoeh, W.R. *et al.* (1991) Heteroplasmy suggests limited biparental inheritance of *Mytilus* mitochondrial DNA. *Science* 251, 1488–1490

3 Zouros, E. *et al.* (1992) Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature* 359, 412–414

4 Yang, X. and Griffiths, A.J. (1993) Male transmission of linear plasmids and mitochondrial DNA in the fungus *Neurospora*. *Genetics* 134, 1055–1062

5 Kondo, R. *et al.* (1992) Further observation of paternal transmission of *Drosophila* mitochondrial DNA by PCR selective amplicification method. *Genet. Res.* 59, 81–84

6 Gyllensten, U. *et al.* (1991) Paternal inheritance of mitochondrial DNA in mice. *Nature* 352, 255–257

7 Hagelberg, E. *et al.* (1999) Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc. R. Soc. London Ser. B* 266, 485–492

8 Eyre-Walker, A. *et al.* (1999) How clonal are human mitochondria? *Proc. R. Soc. London Ser. B* 266, 477–483

9 Awadalla, P. *et al.* (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286, 2524–2525

10 Vigilant, L. *et al.* (1991) African populations and the evolution of human mitochondrial DNA. *Science* 253, 1503–1507

11 Maddison, D.R. *et al.* (1992) Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Syst. Biol.* 41, 111–124

12 Wakeley, J. (1994) Substitution-rate variation among sites and the estimation of transition bias. *Mol. Biol. Evol.* 11, 436–442

13 Hagelberg, E. *et al.* Evidence for mitochondrial DNA recombination in a human population of island Melanesia: correction. *Proc. R. Soc. London Ser. B* (in press)

14 Borresen, A.L. *et al.* (1988) Linkage disequilibrium analyses and restriction mapping of four RFLPs at the pro alpha 2(I) collagen locus: lack of correlation between linkage disequilibrium and physical distance. *Hum. Genet.* 78, 216–221

15 Awadalla, P. and Charlesworth, D. (1999) Recombination and selection at *Brassica* self-incompatibility loci. *Genetics* 152, 413–425

16 Conway, D.J. *et al.* (1999) High recombination rate in natural populations of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* 96, 4506–4511

17 Rich, S.M. *et al.* (1997) *Plasmodium falciparum* antigenic diversity: evidence of clonal population structure. *Proc. Natl. Acad. Sci. U. S. A.* 94, 13040–13045

18 Sawyer, S. (1989) Statistical tests for detecting gene conversion. *Mol. Biol. Evol.* 6, 526–538

19 Maynard Smith, J. and Smith, N.H. (1998) Detecting recombination from gene trees. *Mol. Biol. Evol.* 15, 590–599

20 Luria, S.E. and Delbruck, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28, 491–511

21 Sutovsky, P. and Schatten, G. (2000) Paternal contributions to the mammalian zygote: fertilization after sperm–egg fusion. *Int. Rev. Cytol.* 195, 1–65

22 Rogers, A.R. and Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9, 552–568