DNA variation and evolution

FERRIS, Sage and Wilson[1] assayed wild-caught and inbred strains of mice (*Mus domesticus*) for variation in mitochondrial DNA (mtDNA) using restriction enzyme digestion. They concluded that the data showed either a common origin of 'old' inbred strains (probably a single female) or contamination from a single strain because all the 'old' inbred strains appeared to share a common mtDNA pattern found rarely (2 mice in 56 tested) in their survey of wild populations.

Their observations were made on samples of single mice from each of 14 inbred laboratory strains (5 'new' and 9 'old') and 15 wild populations; further samples of two and three mice were from two different origins. Thirty-six additional wild mice (origin not stated) were assayed using two restriction enzymes which could discriminate the 'old–inbred' pattern. When all members of a sample of size N are identical for a particular pattern, the 95% confidence intervals on the frequency of that pattern are: for $N=1$, 0.025–1.00; for $N=2$, 0.158–1.00; for $N=3$, 0.292–1.00. Because the size of samples used in these experiments was small, it is possible that substantial levels of variation could remain undetected within the wild populations. Whilst the data may suggest a common origin of the 'old' inbred strains, the lack of data on variation within the wild populations precludes the elimination of the alternative hypothesis of independent origin.

In another paper, Coen, Thoday and Dover[2] based their estimation of rates of 'turnover' in structural variants of the ribosomal gene of *Drosophila melanogaster* on samples of two X chromosomes extracted from 10 pairs of iso-female lines and samples of 'at least two (usually three' X chromosomes from population cages established with a mixture of all iso-female lines. The authors assumed that if a variant was present in at least two sublines, it was probably derived from the founding population. On this basis no novel variants were detected in the paired iso-female lines and the rate of 'turnover' was estimated by assuming that no novel variants had been identified. However, no allowance was made for either the size of the samples or the sizes of the populations from which the samples were extracted. The maximum frequency at which novel variants could exist in all sublines and yet remain undetected ($P=0.05$) in samples of 2 from 20 sublines is ~0.07. In other words, the iso-female lines could be segregating for novel variants at fairly high frequencies and these would not be detectable because of the small sample sizes.

If these putative variants have arisen since the establishment of the lines, then the maximum rate of their production (m, the mutation or 'turnover' rate) is about 3.74×10^{-4} per generation which is of the same order as values obtained by Coen *et al*[2]. However, these estimates assume that the populations are large and no allowance is made for the finite size of the subline populations. Downes[3] gives estimates of the effective population sizes (N_e) of the OK iso-female lines of between 90 and 400. At these values of N_e and m, the effects of the random sampling of gametes in each generation could have a considerable effect on the fate of new variants because the majority will be lost in the first generation. Consequently, the chances of detecting 'turnover' would be low in such small populations and the measurement of its rate made more difficult by the small sample sizes. Similarly the small samples taken from the cage populations and absence of any information on the effective population sizes restrict the interpretation of the data. The results quoted by Coen *et al.* provide no reasonable *prima facie* evidence of systematic change in the frequency of rDNA structural variants. The estimates of 'turnover' rates are based on the assumption that 'turnover' is occurring, not on a demonstration of its existence.

The reports of both Ferris *et al.*[1] and Coen *et al.*[2] describe the application of the techniques of molecular biology to investigate evolution. The procedures have exposed a range of variation which evolutionary biologists must consider seriously. However, the fascinating interpretations presented in these papers are not strongly supported by the data because alternative hypotheses cannot be eliminated. Acceptance of the conclusions of the two reports must await the availability of more data.

JOHN A. BARRETT
Department of Genetics,
University of Liverpool,
PO Box 147,
Liverpool L69 3BX, UK

1. Ferris, S. D., Sage, R. D. & Wilson, A. C. *Nature* **295**, 163–165 (1982).
2. Coen, E. S., Thoday, J. M. & Dover, G. A. *Nature* **295**, 564–568 (1982).
3. Downes, S. P. thesis, Univ. Cambridge (1981).

FERRIS *ET AL.* REPLY—'Concluded' is too strong a term to apply to the various possible explanations we *offered* for the remarkable finding[1] that all nine of the 'old inbred' strains of laboratory mice tested share a type of mtDNA whose frequency is low (0.04) in wild mice (*Mus domesticus*). This finding contrasts sharply with published protein-electrophoretic evidence indicating that the nuclear genes of these laboratory strains are about as different from one another as are randomly picked nuclear allelles in wild mice[1]. It was important not only to bring our finding to the attention of a wide audience but also to consider how to explain this puzzling contrast in light of the evidence that the mode of inheritance of mtDNA molecules differs from that of nuclear genes in being maternal.

Since these nine strains seen from published breeding records to be founded from at least five non-laboratory females, we calculated the probability of picking five mice with the 'old inbred' type at random from the wild. By 'at random from the wild', we meant at random geographically within the entire range occupied by wild members of the species *M. domesticus*. The best estimate of this probability, based on the frequency observed in our wild survey, is $(0.04)^5$, that is, 10^{-7}. This value, which has been confirmed by a more extensive survey[2], was so low that it led us to ask whether: (1) the five founding females were picked nonrandomly; (2) there was selection for the old inbred type; (3) there was frequent, unrecorded, contamination of one laboratory strain by another's mtDNA.

It was our position[1], which our subsequent work on many additional mice confirms[2–4], that none of these possible explanations for the mtDNA uniformly of the old inbred strains can yet be ruled out.

With regard to nonrandom sampling, we pointed to published evidence that most or all of the founding females came from the pet mouse trade. Drift or selection within this population could have brought about a high frequency of the 'old inbred' type of mtDNA. Alternatively, most of the wild mice used to establish or maintain the pet mouse population may have been trapped in a geographically nonrandom way, that is, from a particular region in which the 'old inbred' type happened to be at high frequency. Our initial survey of mtDNA in wild mice was a wide-ranging one, designed to detect all the major mtDNA lineages in *M. domesticus*[1]. We were aware that until a much more extensive survey of mtDNA variability within and between wild populations of *M. domesticus* was conducted, this alternative could not be evaluated. Subsequent survey work[2–4] bringing the total number of wild *domesticus* mtDNAs tested to 145, has confirmed our initial estimate of the frequency of the old inbred type; it has also shown, as we anticipated, that while there is heterogeneity within and between some localities as regards