Ann. Hum. Genet. (1996), **60**, 85–86
Printed in Great Britain

Molecular evolution: selection wins the debate

The Causes of Molecular Evolution. By J. H. GILLESPIE. Oxford: Oxford University Press. 1991. pp. 336. H/b (0 19 506883 1) P/b (0 19 50921 6) £14.95

This is an extremely important book. New theories and new molecular data are finally addressing problems that arose when protein variability made classical theories untenable. The neutral theory of molecular evolution, proposed in the 1960s by King, Jukes and Kimura could make broad sweeping predictions simply because it was a null hypothesis that should affect all loci equally. To construct a competing theory of selection, which Gillespie does here, has been far more difficult, requiring the integration of selection variable in both time and space across differently selected loci.

In the book to which Gillespie's is a natural heir, Lewontin (1974) despaired that evolutionary theory could cope with molecular variation: 'The mother lode has been tapped, and facts in profusion have poured into the hoppers of this theory machine. And from the other end has issued – nothing. It is not that the machinery does not work, for a great clashing of gears is clearly audible, if not deafening, but it somehow cannot transform into a finished product the great volume of raw material that has been provided.' Lewontin concluded his negative, although at least unbiased, review of neutral vs. selectionist theories of molecular evolution with an even more negative idea: that most evolutionary questions about single loci were perhaps unanswerable anyway. To Lewontin, most evolution at a locus was caused by hitch-hiking correlations with other loci under selection. Lewontin was wrong; recent studies of the *Alcohol dehydrogenase* locus in *Drosophila melanogaster* suggest linkage disequilibria around a selected site disappear over a few hundred base pairs (Kreitman & Hudson, 1993). In humans, with smaller total populations, disequilibria are likely and are observed over perhaps ten-fold larger physical distances. However, these are still of the order of intragenic rather than intergenic distances. Exceptions to this rule include the MHC loci, thought to be under epistatic selection by pathogens. In order to maintain disequilibria between distant loci, it now seems certain that selection must be both strong and epistatic. These discoveries mean that many, perhaps most, loci can be analysed singly, without having to consider correlations with the rest of the genome.

Those of you who feel enslaved by the human genome mapping project will no doubt be happy to learn that you, together with *Drosophila* geneticists, are providing some of the best data for studies of molecular evolution. There were considerably fewer data available when this book was written in 1990–1991 than are available now. (The book was published in 1991; presumably the reason I have now been given the book for review is that the paperback edition came out in 1994.) A lot of evidence shows that a monolithic neutral theory cannot explain evolution at both silent and replacement sites. Silent sites usually evolve faster, although sometimes slower, than sites which cause amino acid changes. The rate of evolution of both DNA and proteins has a higher variance than mean: the rate of substitution is episodic, rather than Poisson-distributed as expected under the neutral hypothesis. Heterozygote advantage and selection leading to amino-acid substitution is now known to be intense in some individual cases. However, since null enzyme alleles are often relatively common, and since null homozygotes often survive, Gillespie argues that the relevant selection on functional enzyme alleles must often be weak, say $< 0{\cdot}1\%$.

As you might expect, general models of stochastic selection are not simple. Gillespie's idea is that many protein alleles, like the null alleles referred to above, will have 'concave fitnesses', that is they will be dominant with respect to effects on fitness. Temporal variation in direction and strength of selection provides the motor for evolution. Many aspects of this theory are already proved. For instance, the reason for overdominance of the sickle cell allele is that this allele is dominant with respect to its effect on resistance to malaria, and recessive with respect to its effect on severe anaemia. These two types of selection combine to cause heterozygous advantage in the presence of malaria. Gillespie's

explanation of molecular evolution is the same kind of effect writ small, so to speak, on most protein loci, with temporally and spatially varying selection providing both the initiating causes of autocatalytic cascades of substitution and the means to maintain polymorphism.

The book is well produced and clearly, although somewhat gnomically, written. Some of the computer diagrams and graphs could have been made more comprehensible to humans. Important though the book is, Gillespie has written with a mathematical readership in mind, and has not helped the general reader by verbally explaining the aims, assumptions and results of a host of rather reconditely named mathematical models and analyses; examples include the c-haploid model, the SAS-CFF (stochastic additive scale – concave fitness function) model, or the SSWM (strong selection-weak migration) analysis.

Finally, if selection is generally as weak as 10^{-3}, is Gillespie's model testable? Though we have more and better molecular data, which clearly show that replacements are non-independent, it is unclear how well Gillespie's post-hoc model fits the available data and what new predictions it makes that can be tested. But there is hope, because there are so many new molecular patterns we can now investigate. For instance, virtually nothing has been done to compare levels of polymorphism at allozymes and at silent DNA simultaneously. Preliminary evidence suggests that the two types of variation behave very differently in *Drosophila* (Begun & Aquadro, 1993), strongly implicating selection in the maintenance of allozyme polymorphisms. Secondly, closely related species or even populations separated by hybrid zones may differ strongly at some allozyme loci, but remain virtually identical at others, again suggesting strongly non-random effects between loci. As more molecular data are gathered and synthesized from such cases, it seems that we are at last nearing a solution to the causes of molecular evolution. And it is already clear that this solution will mostly involve selection.

References

Lewontin, R. C. (1974) *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Kreitman, M. & Hudson, R. R. (1991) Inferring the evolutionary histories of the Adh and Adh-dup loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. *Genetics* **127**, 565–582.

Begun, D. J. & Aquadro, C. F. (1993) African and North American populations of *Drosophila melanogaster* are very different at the DNA level. *Nature* **365**, 548–550.

J. L. B. Mallet