Letter to the Editor

More on the Overdispersed Molecular Clock

The neutral allele theory in its present form is not able to account for the observed variation in the rate of molecular evolution (GILLESPIE 1987). The purpose of a recent paper by TAKAHATA (1987) was to show that certain modifications of the neutral allele theory would make it compatible with the observed variation. However, when TAKAHATA's modifications are examined in detail, they do not appear to accomplish everything that is claimed.

TAKAHATA quantifies the variation in the rate of molecular evolution as the ratio of the variance in the number of substitutions on a lineage of length t to the mean number of substitutions, $I(t)$. KIMURA's (1983) version of the neutral theory predicts that this ratio should be very close to one. When a large number of loci are examined, $I(t)$ is typically between one and ten, although values as high as 35 have been observed (GILLESPIE 1986). TAKAHATA investigated three extensions of the neutral theory that cause the predicted value of $I(t)$ to be elevated. One of the extensions incorporates fluctuations in the population size, while the others introduce assumptions about the effects of one substitution on the waiting time for subsequent substitutions.

TAKAHATA's first model assumes that substitutions occur in bursts: in a very short period of time, several substitutions take place. In support of this point of view, TAKAHATA cites FITCH and MARKOWITZ's (1970) covarion model. This model assumes that, at any point in time, all of the evolution is restricted to a small subset of the amino acid sites. To understand the problem with this argument, recall that FITCH and MARKOWITZ arrived at the covarion model as an explanation for the "excess" of double substitutions observed in protein sequence data. The excess is only an excess under the assumption that all amino acid sites evolve at the same rate. If certain amino acid sites evolve more rapidly than others, then there will be an *apparent* excess of multiple substitutions at those sites that evolve more rapidly. The covarion model is simply an extreme example of this situation. A more general model would be one in which the rate of evolution is represented as a distribution of rates over sites as described in GILLESPIE (1986). The important point, however, is that these models of rate variation across sites imply *nothing* about the rapidity of successive substitutions; they only tell where they are likely to occur, thus providing no support for the multiple substitution model. TAKA-

HATA himself ultimately rejects the multiple substitution model for quantitative reasons, so the improper motivation does not affect the main thrust of his paper.

The most common explanation for the variability of evolutionary rates is that they are due to fluctuations in the size of the population. To explore this explanation, TAKAHATA assumes that the population fluctuates between two sizes, the larger size characterized by a rate of evolution of $a = 10^{-6}$ amino acid substitutions per loci per year and the smaller size by a rate of $b = 10^{-4}$. The larger size occurs for an average duration of 10^7 years, and the smaller size for 10^5 years. With these assumptions TAKAHATA shows that $I(t) = 3.45$, thus providing apparent support for this explanation. I had made a similar calculation (GILLESPIE 1987) and arrived at the conclusion that fluctuations in the population size would not elevate the variance in the rate of evolution appreciably. The difference in our points of view stem from the assumption each of us made about the mutation rate to nearly neutral alleles. As is well known, the rate of molecular evolution under the mildly deleterious allele model will always be less than the mutation rate to mildly deleterious alleles. Thus in his calculation, TAKAHATA is assuming that the rate of mutation to mildly deleterious alleles is at least 10^{-4} , since this is the only way that a rate of molecular evolution of 10^{-4} could be achieved. Such a rate is unreasonably large, particularly when one considers that mildly deleterious mutations make up only a small fraction of all of the mutations at a locus. If more realistic rates of evolution are used, say 10^{-7} in the larger population and 10^{-6} in the smaller population, then the calculated value of $I(t)$ is 1.04, too small to account for the observed variability of evolutionary rates. The basic problem is that the variance of the rate of evolution is roughly proportional to the mutation rate times the time scale of change of the population size. In order for fluctuations in the population size to elevate the variance in the rate of evolution by a significant amount, the population size must change more slowly (say a change every 10^6 years) than is reasonable. Thus, fluctuations in the population size do not appear able to provide an easy explanation for the elevated values of $I(t)$.

The third model considered by TAKAHATA is the "fluctuating neutral space" model. Under this model

FIGURE 1.—Relationship between relative rate of evolution c and mean number of substitutions $I(t)$.

the rate of evolution changes at random with each substitution. Under the assumption that the rate of evolution is uniformly distributed between a and b , TAKAHATA shows that

$$I(t) = \frac{2}{ab} \left\{ \frac{b-a}{\ln(b/a)} \right\}^2 - 1.$$

This model exhibits a peculiar dependence on the parameters that was not fully described by TAKAHATA. To illustrate this point, let $a = cb$, $0 < c < 1$. Thus, for example, if the lowest rate is 10% less than the highest rate, then $c = 0.9$. The relationship between c and $I(t)$ is illustrated in Figure 1. The dramatic feature of this figure is that for a wide range of values of c , $I(t)$ stays very close to 1, then over a relatively short range, $I(t)$ blasts off to infinity. In order to account for the modal value of $I(t)$ of 2.5 observed in the data, there would have to be very precise restrictions placed on the range of the rate of molecular evolution. While such restrictions may be intrinsic to proteins, there is no evidence for it. Under this model the neutral theory would have no greater predictive value than would any of the many models of evolution based on natural selection. These observations do not depend critically on the uniform distribution of rates. The reader may easily verify

that if rate of evolution is gamma-distributed, then the same extreme parameter dependence will be observed.

Where does this leave us? TAKAHATA acknowledges that the neutral theory in its present form cannot account for the large variance in evolutionary rates. He provides three modifications to the theory to improve the fit, but two of these are inadequate for the reasons mentioned above. The third modification, the "fluctuating neutral space" model, is the most successful, but has no predictive value. More importantly, this model represents such a major departure from the standard neutral model that it is impossible to judge its merits until a significant amount of new work is completed on the implications of this model on the areas where the neutral theory has traditionally been successful in accounting for observations, such as the polymorphism data. The biological basis of this model must also be examined in more detail: it is not entirely clear why the evolutionary rate should suddenly change at the instant that a fixation occurs, but not during the protracted period that the mutant is increasing in frequency prior to its fixation.

JOHN H. GILLESPIE
Department of Genetics
University of California
Davis, California 95616

LITERATURE CITED

- FITCH, W. M., and E. MARKOWITZ, 1970 An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochem. Genet.* **4**: 579-583.
- GILLESPIE, J. H., 1986 Variability of evolutionary rates of DNA. *Genetics* **113**: 1077-1091.
- GILLESPIE, J. H. 1987 Molecular evolution and the neutral allele theory. *Oxford Surv. Evol. Biol.* **4**. In press.
- KIMURA, M., 1983 *The Neutral Allele Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- TAKAHATA, T., 1987 On the overdispersed molecular clock. *Genetics* **116**: 169-179.