Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 8009–8013, December 1984
Population Biology

The molecular clock may be an episodic clock

(molecular evolution/neutral allele theory)

JOHN H. GILLESPIE

Department of Genetics, University of California, Davis, CA 95616

Communicated by G. Ledyard Stebbins, August 22, 1984

ABSTRACT **It is argued that the apparent constancy of the rate of molecular evolution may be an artifact due to the very slow rate of evolution of individual amino acids. A statistical analysis of protein evolution using a stationary point process as the null hypothesis leads to the conclusion that molecular evolution is episodic, with short bursts of rapid evolution followed by long periods of slow evolution. Such dynamics are incompatible with the neutral allele theory and require a revision of the standard interpretation of the molecular clock.**

One of the most enduring generalizations that emerged from early comparative studies of protein sequences is that each protein evolves at a nearly constant rate over lineages (1, 2). This observation has been used as one of the major arguments in support of the neutral allele theory (3), and it is responsible for the concept of a "molecular clock" (2). Although a superficial examination of the protein sequence data does suggest that proteins evolve at a constant rate, more sophisticated statistical analyses have called this generalization into question (4–6). The prevailing point of view appears to be that, although proteins do not evolve at a constant rate, the deviations are not large enough to seriously threaten either neutrality or the molecular clock.

In this paper, it will be shown that the analyses of protein sequence data may have been distorted by unrecognized biases that have led to gross underestimates of the variability in the rates of evolution. If the biases are taken into account, the inferred dynamics of molecular evolution appear to be much more erratic than suggested both by neutral allele models and by the molecular clock hypothesis.

Statistics of Star Phylogenies

The statistical analysis of protein sequence data is extraordinarily difficult. For complex phylogenies, the most severe problem stems from the unknown biases introduced by the algorithms for reconstructing ancestral sequences. However, Kimura (3) recognized that there is one setting in which much can be learned about the moments of the numbers of substitutions that occur on each branch of a tree. This is the case of a "star phylogeny," where the sequences are from species that are derived from a radiation that occurred in a short time relative to the length of the lineages. For such clades, the problems associated with inferring evolutionary trees may be avoided by assuming that all lineages stem simultaneously from a single common ancestor. I begin by briefly describing the estimation procedure developed by Kimura for star phylogenies and will interpret the results in subsequent sections.

Given a star phylogeny, let X_i be the number of substitutions that have occurred on the ith of n lineages stemming from the common ancestor. Whereas a desirable goal would be to achieve a good estimate of the distribution of the X_i, because of the paucity of data we must be satisfied with estimates of the first two moments, assuming that the X_i are identically distributed, but not necessarily independent, random variables. Kimura's (3) method of estimating these moments uses D_{ij}, the number of amino acids that differ between species i and j. The method would be straightforward if D_{ij} equalled $X_i + X_j$. However, because of multiple substitutions at a single site, this equality is violated. Therefore, the first step is to correct the D_{ij} for multiple substitutions and then to proceed as if $D_{ij} = X_i + X_j$. Here, I use Dayhoff's (7) empirically derived acceptable point mutation correction for the D_{ij}. For the data that will be examined, the differences between sequences are so small (typically 10%) that the effect of this correction (or any of the other published corrections) is trivial. Once the corrections are made, the moments are estimated using the new D_{ij} and the three estimation formulas given by Kimura (3). The mean of X_i may be estimated by

$$M = \frac{1}{n(n-1)} \sum_{i<j} D_{ij}.$$

It is easy to see that the expectation of M is

$$E(M) = E(X_i).$$

Similarly, a second order moment of the X_i may be estimated by

$$S^2 = \frac{1}{(n-1)(n-2)} \sum_{i<j} (D_{ij} - 2M)^2.$$

The expectation of S^2 is

$$E(S^2) = \mathrm{Var}(X_i) - \mathrm{Cov}(X_i, X_j).$$

Finally, the index of dispersion, the ratio $\mathrm{Var}(X_i)/E(X_i)$, may be estimated by

$$R = S^2/M.$$

Unlike the previous two estimators, R is biased. The bias decreases as the number of lineages increases. Table 1 gives the values for M, S^2, and R for the five proteins considered by Kimura. The values are slightly different from those in Kimura's book because of the use of the Dayhoff rather than the Zuckerkandl and Pauling (1) correction used by Kimura.

Of particular interest in Table 1 are the estimates of R. It is commonly held that if the rates of evolution are constant and/or if neutrality is the mechanism of evolution, then the X_i will be independent Poisson random variables and, thus, the index of dispersion will be 1. To test this, Kimura argued that, under the Poisson assumption, $(n-1)R$ is χ^2 distributed with $(n-1)$ degrees of freedom and thus provides a convenient test statistic. Of the five proteins examined in the table, two show a significant departure from the Poisson assumption based on this criterion. The significance of this ob-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Table 1. Reanalysis of amino acid sequence data from Kimura (3)

Protein	Number of species, n	Mean substitutions per lineage, M	S^2	R	Mean episodes per lineage, $1/V$	Mean substitutions per episode, MV
Hemoglobin-α	6	13.15	18.30	1.39	67.10	0.20
Hemoglobin-β	6	15.61	54.19	3.47*	12.64	1.24
Myoglobin	6	12.77	23.83	1.87	29.48	0.43
Cytochrome c	4	8.55	30.92	3.62*	6.54	1.31
Ribonuclease	4	21.99	62.68	2.85	23.79	0.93

*These values of R are significantly >1.

servation is by no means obvious. This will be illustrated by developing various point process models of molecular evolution.

Point Process Models of Molecular Evolution

In this and the next section, I concentrate on the events that occur through time in a single lineage. Molecular evolution will be modeled using a stochastic point process. A point process may be thought of as a collection of the points in time when certain events occur. In this case, the events will be the substitution of one amino acid for another in a protein. I am particularly interested in the total number of events (substitutions) that have occurred in an interval of time of duration t. This total will be notated as $N(t)$. $N(t)$ is frequently called a cumulative process. My ultimate goal is to infer certain properties of $N(t)$ using the estimators M, S^2, and R. To ensure that this goal is achievable, various assumptions about $N(t)$ must be accepted. The fundamental assumption is that the point process is stationary. That is, the process is invariant under translations in time. This is a common starting assumption in time series analysis and one that is susceptible to statistical scrutiny. The Poisson process is a special example of a stationary point process. For this process, $N(t)$ is Poisson distributed, implying that the index of dispersion,

$$I(t) = \mathrm{Var}[N(t)]/E[N(t)],$$

is equal to 1. For more general stationary point processes, the index of dispersion is a nonlinear function of t. It will, however, approach 1 for small values of t.

Since a protein is a string of amino acids, there must be a stationary point process associated with each amino acid. Let $Z_i(t)$ represent the cumulative point process for the ith amino acid. Obviously,

$$N(t) = Z_1(t) + Z_2(t) + \cdots + Z_p(t),$$

for a protein composed of p amino acids. A point process that is the sum of other point processes is called a superposed process. If p Poisson processes are superposed, the resulting process will also be a Poisson process. (See ref. 8 for a good exposition of these and other aspects of the theory of point processes used in this paper.)

A problem that must be addressed when developing a statistical model for protein evolution concerns the relationship between the process of change at the level of the individual amino acids [the $Z_i(t)$] and the process at the level of the protein as a whole, $N(t)$. There are two extreme forms for this relationship. The first, and most commonly used, pictures each amino acid as undergoing an independent process of evolution with the evolution of the protein being a simple superposition of these independent processes. In this instance, the index of dispersion of the superposed process will equal a weighted average of the indices of the amino acids. The second pictures the evolution of the separate amino acids as being so highly dependent on one another that the process is best studied at the level of the entire protein without reference to the processes at the level of the individual amino acids. The former point of view has some interesting implications for the interpretation of the data in Table 1.

The Superposition Problem

Viewing the amino acids as undergoing independent point processes implies that the apparent closeness of R to 1 is an artifact of the superposition of the separate processes. This is a consequence of a well-known theorem from queuing theory, the Palm–Khintchine theorem (9), whose essence is captured in the following. Suppose we have p independent, stationary, superposed point processes each of whose rates are proportional to $1/p$. If the superposed process is followed for a fixed period of time, then as p gets large the superposed process will converge to a Poisson process. A simple corollary is that the index of dispersion approaches 1 as the number of superposed processes increases. The theorem is intuitive. Consider that in the case of proteins a typical amino acid is replaced, on the average, about once every 1 billion yr. Thus, the probability of any particular amino acid being replaced in the 10- to 100-million-yr spans of typical studies is small. Since the number of amino acids is large, if the changes are independent, we would expect the distribution of the number of changes in the protein to be approximately Poisson. Thus, even if the index of dispersion of the individual amino acids and, hence, the protein is very large, it will appear to be small for the superposed process because of the restricted period of observation. If lineages could be examined that were 10- to 100-billion-yr long, these problems would be less significant.

To illustrate the effect of superposition, we can simulate on a computer a situation that is similar to the protein sequence data that Kimura analyzed in his book (3). The simulation assumes that the amino acid substitution process is an equilibrium renewal process with logarithmically normally distributed intervals and an index of dispersion equal to 10 for each of the amino acids. The phylogeny is assumed to be a star phylogeny of six species. The parameters of the simulation were chosen to match fairly closely the data for β-hemoglobin. Thus, the process was run long enough for about 10% of the amino acids to have experienced at least one substitution and the maximum number of amino acids was assumed to be 150. To show the effects of superposition, the simulation was run for proteins made up of 1, 15, 50, or 150 amino acids while the rate of evolution for the entire protein was held constant. For each of the 1000 replicates, the value of R was calculated exactly as described in Kimura's book, and Kimura's test for a significant deviation from a Poisson process (at the 5% level) was applied. Table 2 shows that the results are striking: for 150 amino acids, the null hypothesis that $R = 1$ would be rejected only 8% of the time even though the actual index of dispersion is 10; for 50 amino acids, it would be rejected 22% of the time. This simu-

Table 2. Simulation results for a star phylogeny of six species, where the amino acid substitution processes are independent renewal processes with an index of dispersion equal to 10

Number of amino acids	Average R	Fraction of rejections of H_0
1	4.88	0.831
15	2.07	0.358
50	1.60	0.223
150	1.17	0.084

lation is intended only to point out the incredible bias that is inherent in this type of analysis. Other processes would give different numeric results. Nonetheless, it would be interesting to get some idea of the index of dispersion that is necessary to account for the value of R (3.4) that is observed for β-hemoglobin. Simulations suggest that the actual value could be as high as 1000! Thus, the claim made by Kimura that the observed range of values of R argue for indices of dispersion close to 1 and, hence, neutrality, is vacuous. What is remarkable is that in 40% of the proteins the Poisson process could be rejected at all. This would seem to imply that the amino acid processes are not independent and that the process should be studied at the level of the entire protein. This will be attempted in the next section.

A Model of a Changing Clock

A natural extension to the Poisson process model would be a doubly stochastic Poisson process. This is a Poisson process for which the rate of the process is itself a stationary stochastic process. Not only is this a very general class of point processes, it also admits the interpretation of being a model of a molecular clock with a randomly changing tick rate. Because of the arguments of the previous section, it is assumed that the clock applies to the entire protein rather than to the individual amino acids.

For a doubly stochastic Poisson process, the probability that a substitution occurs in a small interval of time $(t, t + \tau)$ in the ith lineage is $\lambda\theta_i(t)\tau$, while the probability of more than one substitution occurring is of the order of τ^2. In this expression, $\theta_i(t)$ is assumed to be a stationary stochastic process bounded below by 0 with mean $E[\theta_i(t)] = 1$ and autocovariance function $\sigma^2\rho(x)$. $\lambda\theta_i(t)$ will be called the tick rate of the clock. For the ith lineage, let $N_i(t)$ be the number of events that occurred in the interval $(0, t)$. Thus $N_i(t)$ is equivalent to the random variable X_i described in the section on the statistics of star phylogenies.

To complete the description of the model, we must specify the relationships between the clocks in the different lineages of the star phylogeny. One possible assumption is that the tick rate of the clock is assigned independently for each lineage and that the rate is held constant throughout the lineage [$\rho(x) \equiv 1$]. While such a model would fit the data, it is difficult to see why the rate of evolution would change only at the origin of the lineage that is under study and then remain fixed, even though other lineages (not under study) are branching off. A more realistic assumption gives the clocks equal tick rates at the time of the radiation and has the correlation in the tick rates between lineages drop off at the same rate as the correlation within a lineage. In accepting this model, we assume that at the moment of the radiation into the n lineages (at $t = 0$), the processes $\theta_i(t)$ will all equal the same (random) value [i.e., $\theta_i(0) = \theta_j(0)$] and from then on will change independently. Thus, any correlations in the rates of evolution in the separate lineages will be attributable to similar environments at the split times. The correlations will persist for a significant portion of the lineage only if the tick rates of the clocks change very slowly.

Given a particular trajectory of $\theta_i(t)$, $N_i(t)$ is Poisson distributed with mean

$$\lambda I_i(t) = \lambda \int_0^t \theta_i(x)dx.$$

The unconditional moments of the distribution of the number of substitutions depend on the first two moments of the integral $I_i(t)$. Using standard arguments (8), these are

$$E[I_i(t)] = t,$$

$$\mathrm{Var}[I_i(t)] = 2\sigma^2 \int_0^t (t - x)\rho(x)dx = 2\sigma^2 v(t),$$

$$\mathrm{Cov}[I_i(t), I_j(t)] = \sigma^2 \int_0^{2t} (t - |t - x|)\rho(x)dx = \sigma^2 c(t),$$

in which we have taken the opportunity to define the functions $v(t)$ and $c(t)$. The distribution of the number of substitutions in the ith lineage is a randomized Poisson distribution with moments

$$E[X_i(t)] = \lambda t,$$

$$\mathrm{Var}[X_i(t)] = \lambda t + \lambda^2 2\sigma^2 v(t),$$

$$\mathrm{Cov}[X_i(t), X_j(t)] = \lambda^2\sigma^2 c(t).$$

The description of the statistical model being complete, we are now in a position to interpret the data in Table 1. Using the above results, the expectations of M and S^2 become

$$E(M) = \lambda t,$$

$$E(S^2) = \lambda t + \lambda^2 2\sigma^2[v(t) - 1/2\ c(t)].$$

This suggests that the quantity

$$V = (S^2 - M)/(2M^2)$$

might be a good estimator for the variance in the clock, because

$$V \to \sigma^2\psi(t, \rho) \text{ as } n \to \infty,$$

where ψ depends only on the autocorrelation of the process and t. Although this is a consistent estimator, it is, of course, a biased estimator. At the present time, nothing is known about the extent of the bias.

We now turn to the interpretation of V in terms of the second-order moments of the processes $\theta_i(t)$. It is clear from the analysis that the variance in the clock, σ^2, cannot be separated from the autocorrelation as captured in ψ. It is of interest, therefore, to explore the behavior of ψ under various assumptions about the autocorrelation to see how ψ affects the estimate of σ^2. As a simple example, we might suppose that the clock is characterized by an autocovariance function of the form $\sigma^2 \exp(-\alpha|t|)$, which is typical for many Markov processes, such as the Ornstein–Uhlenbeck process and various jump processes. For this clock,

$$\psi = (1/\alpha t)[1 - (1/\alpha t)[1 - \exp(-\alpha t)] \{1 + 1/2[1 - \exp(-\alpha t)]\}].$$

This function depends on t only through the product αt. As αt approaches both 0 and ∞, ψ approaches 0. The maximal value that ψ attains is $\approx$0.19. Thus, the variance in the clock

is >5 times higher than V and could be very large if αt is either very small or very large. Let us consider the meaning of these two extremes. α is interpretable as the rate of decrease in the autocorrelation with time. Thus, if $(1/\alpha) << t$, the time scale of the variation in the tick rate is much shorter than the length of the lineage. On the other hand, if $(1/\alpha) >> t$, the correlations in the tick rate extend beyond the length of the lineages—in our cases, this would be for time spans of tens to hundreds of millions of years. Both of these cases require very large variances in the clock to account for the observed values of V, but for very different reasons. In the former case, the clock is changing rapidly with respect to the sampling interval (the length of the lineage), so the substitution process is made less variable by the smoothing effect of the integral of the clock. Such a clock will be called a rapidly changing clock. In the latter case, the high estimated variance stems from the fact that the $N_i(t)$ are highly correlated with one another so that the variation between the $N_i(t)$ as inferred from the D_{ij} captures less of the variance then it would were the $N_i(t)$ independent. A clock with this behavior will be termed a slowly changing clock. The large variance implied by both of these extremes imposes a very definite behavior on the clock: since $\theta_i(t)$ must be non-negative and have a mean equal to 1, it must spend most of its time near 0 and make only occasional excursions to very large values. For a rapidly changing clock, these excursions will be of very short duration relative to the length of the lineage, suggesting that this clock be called an episodic clock.

This treatment can be extended to general stationary processes. For rapidly changing clocks we have

$$V \sim \sigma^2 \int_0^\infty \rho(x)dx/t,$$

where this approximation assumes that $\int\rho(x)dx/t << 0$. For slowly changing clocks we have

$$V \sim \sigma^2 (\rho_1 t/3 + \rho_2 t/8),$$

where we have used the first terms of the Taylor series expansions of $\rho(x)$:

$$\rho(x) \sim 1 - \rho_1 x - (\rho_2/2)x^2,$$

and assumed that $\rho_1 t$ and $\rho_2 t^2$ are $<<1$. Notice that in this more general setting, the inferred variance in the clock also increases with both decreasing and increasing autocorrelation.

Implications About Molecular Evolution

The confounding of the variance with the autocorrelation of the clock presents a dilemma for the interpretation of molecular evolution data. Is the slowly changing or the episodic clock more appropriate? As there is no biological connection between the autocorrelation of the clock and the lengths of lineages used in molecular evolution studies, it seems unlikely that the time scale of the clock will equal the length of the lineages. This makes it unlikely that a clock intermediate between the two extremes will be relevant. Since the lineages that are commonly used in molecular evolution studies have common ancestors in the late Mesozoic or earlier, the time scale of the slowly changing clock would have to be longer than ≈66 million yr. Such long time scales of change seem incompatible with the fact that many major environmental changes, such as the recent ice ages, occur on time scales of tens of thousands to hundreds of thousands of years. Thus, of the two clocks, the episodic clock seems more in accord with our usual ideas of the time scales of environmental change. In addition, if the slowly changing clock were correct, one would expect a very different pattern in the number of substitutions per lineage than is actually observed. Namely, most of the lineages would be expected to have no substitutions at all and only an occasional lineage would be expected to have a relatively large number of substitutions. This pattern is never observed in the data. By contrast, under the episodic clock, the episodes should be sprinkled at random among the lineages and, consequently, the number of substitutions would be more equally dispersed. Or, said another way, for the same mean and variance the slowly changing clock will produce a much more skewed distribution of the number of substitutions per lineage than is observed. For these reasons, the remainder of this paper is concerned with properties of the episodic clock.

A simple version of the episodic clock is a two-state Markov jump process that takes on the value 0 for an exponentially distributed length of time with mean $1/\mu_0$ before jumping to the value β, where it remains for an exponentially distributed length of time with mean $1/\mu_1$ before returning to 0. For this version of $\theta_i(t)$ to have a mean equal to 1, we require that $\beta = \mu_1/\mu_0$. The requirement that the time scale be much less than the length of the lineage and that the variance is large is met if we assume that $\mu_1 \to \infty$. As this limit is approached, the autocovariance function approaches

$$(\mu_1/\mu_0) \exp(-\mu_1|t|),$$

and the mean length of the episodes of rapid evolution decreases. This allows the integral of θ to be approximated by the sum of a Poisson number of random variables, Z_i, each equaling the product of β $(=\mu_1/\mu_0)$ and an independent exponentially distributed random variable with mean $1/\mu_1$. This product is simply the area under $\theta_i(t)$ for the ith episode. The mean of the Poisson distribution approaches $\mu_0 t$. During each episode, the actual number of substitutions will also be Poisson distributed with mean $Y_i = \lambda Z_i$ conditioned on the value of Z_i. Thus, we can approximate the distribution of the total number of substitutions on a lineage as the sum

$$Y = Y_1 + Y_2 + \cdots + Y_{M(t)},$$

where

$$\begin{aligned} E(Y_i) &= \lambda/\mu_0, \\ \mathrm{Var}(Y_i) &= (\lambda/\mu_0)(1 + \lambda/\mu_0), \\ E[M(t)] &= \mu_0 t, \\ E(Y) &= \lambda t, \end{aligned}$$

and

$$\mathrm{Var}(Y) = \lambda t + 2(\lambda t)^2/(\mu_0 t).$$

Thus, our estimator V from a previous section estimates $1/\mu_0 t$, suggesting that $1/V$ may be interpreted as the mean number of episodes per lineage. The mean number of substitutions per episode is the mean number of substitutions per lineage divided by the mean number of episodes per lineage. Finally, the ratio of the variance in the number of substitutions to the mean number for each episode is equal to the ratio for the total number of substitutions under this approximation. All of this is summarized in Table 1. It should be noted that the distribution of the number of substitutions per protein under this model is a randomized Poisson distribution with a variance larger than the mean. This is in accord with the recent demonstration that the spatial pattern of substitutions on the molecule fits the negative binomial distribution better than the Poisson distribution (10).

It may be possible to extend this analysis to more general models of the tick rate than the two-state Markov process. The representation of the integral of the clock as a Poisson sum of independent random variables may be generally valid for any tick-rate process, as similar representations may be found in the theory of Z_n exceedance measures for stationary stochastic processes (11). All of this strongly suggests that the appropriate null model for molecular evolution would be a cluster process of the form

$$N(t) = Y_1 + Y_2 + \cdots + Y_{M(t)},$$

where the Y_i are independent identically distributed random variables, and $M(t)$ is a stationary point process. If the doubly stochastic Poisson process version of the clock is viewed as a reasonable guide, then $M(t)$ may be assumed to be a Poisson process. Otherwise, and more cautiously, we should view $M(t)$ as having some correlation structure.

These results have some bearing on the use of protein sequence data to date the split times of lineages. They suggest that the estimates of the variance in time back to a common ancestor could be gross underestimates unless there is a calibration using sequences from species with known split times that are similar to the one under study.

Implications for the Mechanism of Molecular Evolution

One of the outstanding questions in molecular evolution concerns the mechanisms responsible for the amino acid differences that occur between species. An attractive theory for these differences is the neutral allele theory (3). However, our statistical analysis suggests that the course of molecular evolution is episodic, being quite unlike the dynamics of the neutral allele theory. This should not be surprising, given Hudson's (12) recent demonstration that the neutral allele model with a constant mutation rate is not compatible with the available data for protein differences both within and between species. One resolution to this problem might be to assume that the mutation rates are themselves episodic in nature. The consequences of this may well increase the variance in the number of substitutions per lineage while lowering the mean heterozygosity.

Alternatively, the episodic clock may be due to the action of natural selection. As an example, I (13) recently described the effects of the mutational structure of DNA on the variance to mean ratio of the number of substitutions per lineage when natural selection is responding to continuing changes in the environment. The conclusion was that the structure of DNA will cause evolution to proceed in a series of bursts, even if the changes in the environment proceed in a Poisson fashion. The burst structure of that model closely resembles the episodic character of the clock. Another possibility is that the episodes of rapid evolution are associated with periodic environmental changes, speciation events, or crises of extraterrestrial origin. However, such speculations are premature.

I thank James F. Crow, Joseph Felsenstein, Robin Gordon, and Michael Turelli for numerous suggestions that improved the final version of this paper. This acknowledgement does not imply that these individuals necessarily agree with all aspects of the conclusions of this paper. The simulations were performed using an 8087 integrated circuit kindly provided by the Intel Corporation.

1. Zuckerkandl, E. & Pauling, L. (1965) in *Evolving Genes and Proteins*, eds. Bryson, V. & Vogel, H. J. (Academic, New York), pp. 97–166.
2. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639.
3. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, England).
4. Ohta, T. & Kimura, M. (1971) *J. Mol. Evol.* **1**, 18–25.
5. Langley, C. H. & Fitch, W. M. (1974) *J. Mol. Evol.* **3**, 161–177.
6. Fitch, W. M. & Langley, C. H. (1976) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **35**, 2092–2097.
7. Dayhoff, M. O., ed. (1972) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3.
8. Cox, D. R. & Isham, V. (1980) *Point Processes* (Methuen, London).
9. Khintchine, A. Y. (1960) *Mathematical Methods in the Theory of Queueing* (Griffin, London).
10. Holmquist, R., Goodman, M., Conroy, T. & Czelusniak, J. (1983) *J. Mol. Evol.* **19**, 437–448.
11. Cramer, H. & Leadbetter, M. R. (1967) *Stationary and Related Stochastic Processes* (Wiley, New York).
12. Hudson, R. R. (1983) *Evolution* **37**, 203–217.
13. Gillespie, J. H. (1984) *Evolution*, in press.