Proc. Natl. Acad. Sci. USA
Vol. 76, No. 7, pp. 3440–3444, July 1979
Genetics

Model of effectively neutral mutations in which selective constraint is incorporated

(molecular evolution/protein polymorphism/population genetics/neutral mutation theory)

MOTOO KIMURA

National Institute of Genetics, Mishima 411, Japan

Contributed by Motoo Kimura, April 25, 1979

ABSTRACT **Based on the idea that selective neutrality is the limit when the selective disadvantage becomes indefinitely small, a model of neutral (and nearly neutral) mutations is proposed that assumes that the selection coefficient (s') against the mutant at various sites within a cistron (gene) follows a Γ distribution; $f(s') = \alpha^\beta e^{-\alpha s'} s'^{\beta-1}/\Gamma(\beta)$, in which $\alpha = \beta/\bar{s}'$ and $\bar{s}'$ is the mean selection coefficient against the mutants ($\bar{s}' > 0$; $1 \geqq \beta > 0$). The mutation rate for alleles whose selection coefficients s' lie in the range between 0 and $1/(2N_e)$, in which N_e is the effective population size, is termed the effectively neutral mutation rate (denoted by v_e). Using the model of "infinite sites" in population genetics, formulas are derived giving the average heterozygosity ($\bar{h}_e$) and evolutionary rate per generation (k_g) in terms of mutant substitutions. It is shown that, with parameter values such as $\beta = 0.5$ and $\bar{s}' = 0.001$, the average heterozygosity increases much more slowly as N_e increases, compared with the case in which a constant fraction of mutations are neutral. Furthermore, the rate of evolution per year (k_1) becomes constant among various organisms, if the generation span (g) in years is inversely proportional to $\sqrt{N_e}$ among them and if the mutation rate per generation is constant. Also, it is shown that we have roughly $k_g = v_e$. The situation becomes quite different if slightly advantageous mutations occur at a constant rate independent of environmental conditions. In this case, the evolutionary rate can become enormously higher in a species with a very large population size than in a species with a small population size, contrary to the observed pattern of evolution at the molecular level.**

Among difficult questions that confront the neutral mutation theory purporting to treat quantitatively the evolution and variation at the molecular level, the following two are particularly acute. First, why the evolutionary rate in terms of mutant substitutions is roughly constant per year for each protein (such as hemoglobin α; see refs. 1 and 2) among diverse lineages, even if the mutation rate appears to be constant per generation rather than per year. Secondary, why the observed level of the average heterozygosity stays mostly in a rather narrow range (between 0% and 20%; see ref. 3) among various species, even if their population sizes differ enormously.

The present paper proposes a model of neutral mutations in which selective constraint (negative selection) is incorporated, and shows that the model can go a long way toward solving these problems in the framework of the neutral mutation theory (4, 5). The model is based on the idea that selective neutrality is the limit when the selective disadvantage becomes indefinitely small (2). For the mathematical formulation of this idea, we must consider the distribution of the selection coefficients of new mutations at the neighborhood of strict neutrality (6, 7). Recently, Ohta (8) investigated a model in which the selection coefficients against the mutants follow an exponential distribution. From the standpoint of the neutral mutation theory, however, Ohta's model has a drawback in that it cannot accommodate enough mutations that behave effectively as neutral when the population size gets large. This difficulty can be overcome by assuming that the selection coefficients follow a Γ distribution.

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.

MODEL OF EFFECTIVELY NEUTRAL MUTATIONS

Let us assume that the frequency distribution of the selective disadvantage (denoted by s') of mutants among different sites within a gene (cistron) follows the Γ distribution

$$f(s') = \alpha^\beta e^{-\alpha s'} s'^{\beta-1}/\Gamma(\beta), \qquad [1]$$

in which $\alpha = \beta/\bar{s}'$, $\bar{s}'$ is the mean selective disadvantage, and β is a parameter such that $0 < \beta \leqq 1$. If we measure the selective advantage in terms of Fisher's Malthusian parameter (9), s' has the range $(0, \infty)$. On the other hand, if we measure it, as we shall do in this paper, in terms of conventional selection coefficient, the true range of s' is restricted to the interval (0, 1). However, because we assume that $\bar{s}'$ is small, with a typical value of 10^{-3}, $f(s')$ is negligible beyond $s' = 0.1$ so that we can take the entire positive axis as the range of integration without serious error. Note that in this formulation, we disregard beneficial mutants, and restrict our consideration only to deleterious and neutral mutations. Admittedly, this is an oversimplification, but as I shall show later, a model assuming that beneficial mutations also arise at a constant rate independent of environmental changes leads to unrealistic results.

Let us consider a diploid population of the effective size N_e and denote by v_e the effectively neutral mutation rate that is defined by the relationship

$$v_e = v \int_0^{1/(2N_e)} f(s')\mathrm{d}s', \qquad [2]$$

in which v is the total mutation rate. For $2N_e\bar{s}' \gg 1$, Eq. 2 is approximated by

$$v_e = \frac{v}{\Gamma(1+\beta)} \left(\frac{\beta}{2N_e\bar{s}'}\right)^\beta. \qquad [3]$$

Fig. 1 illustrates the distribution $f(s')$ for the case $\beta = 0.5$ and $\bar{s}' = 10^{-3}$. In this figure, the shaded area represents the fraction of effectively neutral mutations (v_e/v) when the effective population size (N_e) is 2500. This fraction becomes smaller as the population size increases. Note that even if the frequency of strictly neutral mutations (for which $s' = 0$) is zero in the present model, a large fraction of mutations can be effectively neutral if β is small [note that $f(0) = \infty$ for $0 < \beta < 1$]. We may regard β as representing the degree of physiological homeostasis, while $\bar{s}'$ represents the degree of functional constraint of the molecule. In the limiting situation $\beta \to 0$, all mutations become neutral. On the other hand, if $\beta = 1$, the model reduces to Ohta's model (8) for which $v_e/v \approx 1/(2N_e\bar{s}')$ when $2N_e\bar{s}' \gg 1$.

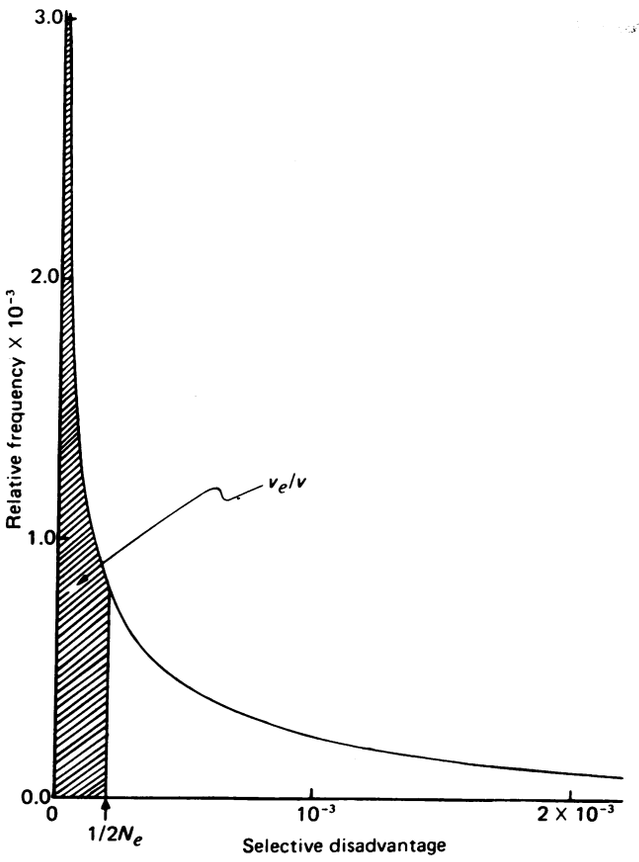

FIG. 1. Frequency distribution of selection coefficients among mutants at different sites within a cistron (gene). The shaded area represents the fraction of effectively neutral mutations. Parameter values assumed are $\beta = 0.5$ and $\bar{s}' = 0.001$. For details, see text.

EVOLUTIONARY RATE

In order to calculate the rate of evolution in terms of mutant substitutions, we assume that the number of available sites (nucleotide or codon sites) for mutation is sufficiently large, while the mutation rate per site is very low so that whenever a mutation occurs it represents a new site in which no mutant forms are segregating within the population. This assumption is known as the model of infinite sites in population genetics. This model was originally formulated (10) with all the nucleotide sites of the genome in mind. The number of nucleotide sites making up a single gene is much smaller, being of the order of several hundreds. Nevertheless, we may apply the infinite site model to a gene locus as a reasonable approximation if the number of segregating sites per gene constitutes a small fraction. It is known (10) that under this model if v is the total mutation rate and if all the mutations are neutral, the expected number of segregating sites is

$$I_1 = 4N_e v[\log_e(2N) + 1], \qquad [4]$$

in which N and N_e are, respectively, the actual (apparent) and the effective sizes of the population. If the mutations are deleterious, the number of segregating sites is smaller. So, I_1 in Eq. 4 may be used to check if the infinite site model is appropriate. As a typical situation, we take $v = 2 \times 10^{-6}$, $N_e = 10^5$, and $N = 10^6$; then we get $I_1 = 12.4$. This constitutes a small fraction compared with several hundred, so the infinite site model may be applicable. However, as N_e becomes larger, I_1 soon gets large enough so that the assumption of infinite sites becomes no longer valid. In such cases, the treatment gives overestimates particularly for the level of heterozygosity. The treatment may still be useful to obtain the upper limit to the heterozygosity.

With this precaution, we proceed to calculate the rate of evolution in terms of mutant substitutions. Let us assume that the mutant is semidominant in fitness so that selection coefficients against the mutant homo- and heterozygotes at a site are $2s'$ and s', respectively. For such a mutant, the probability of eventual fixation in the population is given by

$$u = [1 - e^{2s'(N_e/N)}]/(1 - e^{4N_e s'}), \qquad [5]$$

or, if s' is small,

$$u = 2s'(N_e/N)/(e^{4N_e s'} - 1) \qquad [6]$$

to a good approximation. For the rationale of Eq. 5 see ref. 11, particularly p. 426. Then the rate of mutant substitution per generation is

$$k_g = \int_0^\infty 2Nvuf(s')ds', \qquad [7]$$

in which the subscript g denotes that it refers to the rate per generation rather than per year. Eq. 7 is based on the consideration that the expected number of new mutants that arise in the population in each generation having selective disadvantage in the range $s' \sim s' + ds'$ is $2Nvf(s')ds'$, of which the fraction u eventually reaches fixation in the population. Substituting Eqs. 1 and 6 in Eq. 7, we get, after some computation,

$$k_g = v\beta R^\beta \sum_{j=0}^{\infty} (j + 1 + R)^{-\beta-1}, \qquad [8]$$

in which $R = \beta/(4N_e\bar{s}')$. In Fig. 2, k_g is shown by a solid curve taking the effective population size (N_e) as the abscissa, and

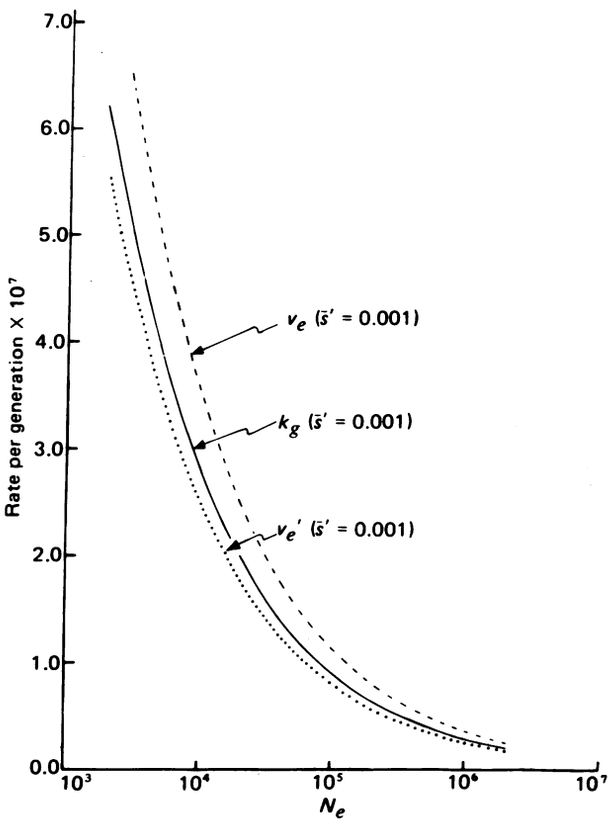

FIG. 2. Comparison between the evolutionary rate (k_g; ——) and the effectively neutral mutation rate (v_e; - - -). ·····, Mutation rate v_e', the rate of occurrence of mutations whose selective disadvantage is less than $1/(4N_e)$. For all curves, $\beta = 0.5$.

assuming $v = 2 \times 10^{-6}$, $\beta = 0.5$, and $\bar{s}' = 0.001$. In the same figure, the effectively neutral mutation rate v_e is plotted by a broken curve for the same set of parameters. Because

$$\sum_{j=0}^{\infty} (j + 1 + R)^{-\beta-1} \approx (1 + R)^{-\beta-1} + \int_{1.5}^{\infty} (\lambda + R)^{-\beta-1} d\lambda = (1 + R)^{-\beta-1} + \beta^{-1}(1.5 + R)^{-\beta}, \quad [9]$$

Eq. 7 may be approximated by

$$k_g = vR^{\beta}[\beta \cdot (1 + R)^{-\beta-1} + (1.5 + R)^{-\beta}]. \quad [10]$$

This approximation gives about 17% overestimation for $\beta = 0.5$ and $N_e\bar{s}' \geqq 1$, but it is accurate enough for most practical purposes. From this we can easily show that, at the limit of either $4N_e\bar{s}' \to 0$ or $\beta \to 0$, we get

$$k_g = v, \quad [11]$$

which is a well-known result for strictly neutral mutations (4). We can also show that, for $4N_e\bar{s}' \gg 1$,

$$k_g \approx v[\beta^{\beta+1}/2^{\beta} + (\beta/3)^{\beta}]/(2N_e\bar{s}')^{\beta}. \quad [12]$$

Comparison of this with Eq. 3 suggests that, roughly speaking, we have

$$k_g \approx v_e. \quad [13]$$

Rough agreement of k_g (solid curve) and v_e (broken curve) may be seen in Fig. 2, in which v_e' (the rate of occurrence of mutations whose s' value is less than $1/4N_e$) is also plotted by a dotted curve for the same set of parameters as used for the other two curves. Thus, Eq. 13 may be regarded as an extension of Eq. 11. In the case of $\beta = 0.5$ as illustrated, the rate of evolution per generation is inversely proportional to $\sqrt{N_e}$ when $N_e\bar{s}'$ is large.

MEAN HETEROZYGOSITY

Let H_n be the expected number of heterozygous sites. Then, as shown in ref. 10 (see equation 15′ therein), we have, for a given value of s',

$$H_n = \frac{8Nv}{-2s'}\left(u - \frac{1}{2N}\right), \quad [14]$$

in which u is given by Eq. 6. This can also be expressed as

$$H_n = 8N_e v \frac{e^{4N_e s'} - 1 - 4N_e s'}{4N_e s'(e^{4N_e s'} - 1)} = 8N_e v \sum_{i=1}^{\infty} \frac{(4N_e s')^i}{(i + 1)!} \sum_{j=1}^{\infty} e^{-4N_e s' j}. \quad [15]$$

Thus the mean number of heterozygous sites, when s' follows the Γ distribution 1, is

$$\overline{H}_n = \int_0^{\infty} H_n f(s') ds' = 8N_e v R^{\beta} \sum_{i=1}^{\infty} \frac{\Gamma(i + \beta)}{\Gamma(\beta)(i + 1)!} \sum_{j=1}^{\infty} (j + R)^{-i-\beta}, \quad [16]$$

in which $R = \beta/(4N_e\bar{s}')$. Then, if we introduce the approximation

$$\sum_{j=1}^{\infty} (j + R)^{-i-\beta} \approx (1 + R)^{-i-\beta} + \int_{1.5}^{\infty} (\lambda + R)^{-i-\beta} d\lambda = (1 + R)^{-i-\beta} + (1.5 + R)^{1-i-\beta}/(i + \beta - 1),$$

Eq. 16 becomes

$$\overline{H}_n = 8N_e v R^{\beta}\left[\sum_{i=1}^{\infty} \frac{\Gamma(i + \beta)}{\Gamma(\beta)(i + 1)!}(1 + R)^{-i-\beta} + \sum_{i=1}^{\infty} \frac{\Gamma(i + \beta - 1)}{\Gamma(\beta)(i + 1)!}(1.5 + R)^{1-i-\beta}\right]. \quad [17]$$

In the special case of $\beta = 1$, this reduces to

$$\overline{H}_n = \frac{8N_e v}{\bar{S}'}\left\{\frac{1}{1 + \bar{S}'} + \log_e(1 + \bar{S}') + \left(\frac{1}{2} + \frac{1}{\bar{S}'}\right)\log_e \frac{1 + 0.5\bar{S}'}{1 + 1.5\bar{S}'}\right\}, \quad [18]$$

in which $\bar{S}' = 4N_e\bar{s}'$.

In order to obtain a simpler expression for Eq. 17 for the case $0 < \beta < 1$, we start from the following formal expression.

$$\sum_{n=0}^{\infty} \frac{\alpha(\alpha - 1)\ldots(\alpha - n + 1)}{n!} x^n = (1 + x)^{\alpha}. \quad [19]$$

Letting $\alpha = -\beta$ and substituting $-xt$ for x in this formula, and then integrating both sides of the resulting equation with respect to t over the interval (0, 1), we get

$$\sum_{n=0}^{\infty} \frac{\beta(\beta + 1)\ldots(\beta + n - 1)}{(n + 1)!} x^{n+1} = \frac{1 - (1 - x)^{1-\beta}}{1 - \beta}. \quad [20]$$

Next, substituting xt for x in Eq. 20, and integrating both sides of the resulting equation with respect to t over the interval (0, 1), followed by putting $n = i - 1$, we have

$$\sum_{i=1}^{\infty} \frac{\Gamma(i + \beta - 1)}{\Gamma(\beta)(i + 1)!} x^{i+1} = \frac{x}{1 - \beta} - \frac{1 - (1 - x)^{2-\beta}}{(1 - \beta)(2 - \beta)}.$$

Finally, if we substitute x^{-1} for x in this equation and then multiply $x^{2-\beta}$ through both sides we get

$$\sum_{i=1}^{\infty} \frac{\Gamma(i + \beta - 1)}{\Gamma(\beta)(i + 1)!} x^{1-i-\beta} = \frac{x^{1-\beta}}{1 - \beta} - \frac{x^{2-\beta} - (x - 1)^{2-\beta}}{(1 - \beta)(2 - \beta)}. \quad [21]$$

Going back to Eq. 20, we note that with a slight modification this may be expressed as

$$\sum_{i=0}^{\infty} \frac{\Gamma(i + \beta)}{\Gamma(\beta)(i + 1)!} x^i = \frac{x^{-1} - x^{-1}(1 - x)^{1-\beta}}{1 - \beta}.$$

Substituting x^{-1} for x in this equation, and after some rearrangements, we get

$$\sum_{i=1}^{\infty} \frac{\Gamma(i + \beta)}{\Gamma(\beta)(i + 1)!} x^{-i-\beta} = \frac{x^{1-\beta} - (x - 1)^{1-\beta}}{1 - \beta} - x^{-\beta}. \quad [22]$$

Applying Eq. 22 with $x = 1 + R$ and Eq. 21 with $x = 1.5 + R$ to the right hand side of Eq. 17, we obtain

$$\overline{H}_n = 8N_e v R^{\beta}\left[\frac{(1 + R)^{1-\beta} + (1.5 + R)^{1-\beta} - R^{1-\beta}}{1 - \beta} - (1 + R)^{-\beta} - \frac{(1.5 + R)^{2-\beta} - (0.5 + R)^{2-\beta}}{(1 - \beta)(2 - \beta)}\right], \quad [23]$$

in which $R = \beta/(4N_e\bar{s}')$. In the limiting situation either for $\beta \to 0$ or $\bar{s}' \to 0$, this equation reduces to $\overline{H}_n = 4N_e v$, which agrees with the result obtained in ref. 10 for strictly neutral mutations.

Let $\overline{h}_e$ be the expected heterozygosity of the gene under consideration. Then, assuming that different sites behave independently, we have

$$\overline{h}_e = 1 - e^{-\overline{H}_n}, \quad [24]$$

because this represents the probability that the gene is heterozygous at least in one of the sites. In Fig. 3, the expected heterozygosity is shown as a function of N_e for various parameter values (β and $\bar{s}'$) assuming $v = 2 \times 10^{-6}$. Note that as compared with the situation in which all the mutations are neutral ($\beta \to 0$), a case such as $\beta = 0.5$ and $\bar{s}' = 0.001$ is interesting because

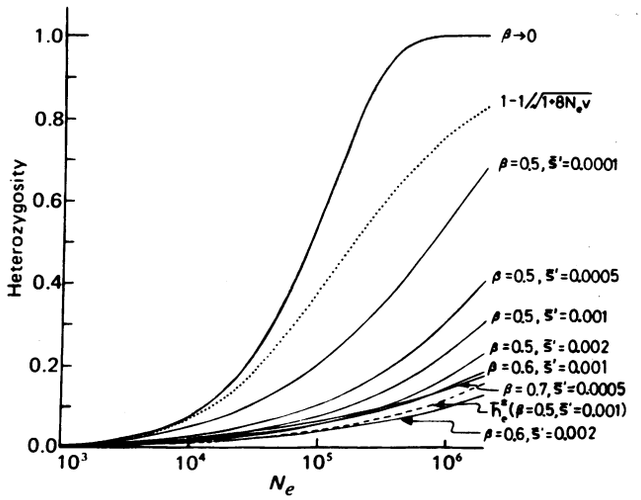

FIG. 3. The mean heterozygosity as a function of the effective population size for various combinations of parameters, β and $\bar{s}'$, assuming the mutation rate $v = 2 \times 10^{-6}$.

of slow rate at which heterozygosity increases as N_e gets large. In the same figure, the dotted line represents $\bar{h}_e = 1 - 1/(1 + 8N_e v)^{1/2}$, the heterozygosity expected under the stepwise mutation model (12). Similarly, the broken line represents $\bar{h}_e^* = 1 - 1/(1 + 8N_e v_e)^{1/2}$ in which v_e is the effectively neutral mutation rate with $\beta = 0.5$ and $\bar{s}' = 0.001$. In all these cases the mutation rate $v = 2 \times 10^{-6}$ is assumed.

SLIGHTLY ADVANTAGEOUS MUTATIONS

To make our analysis complete, let us investigate how the evolutionary rate is influenced by assuming that a certain fraction of mutations are advantageous. Let v_a be the rate of occurrence of advantageous mutations, and assume that the selection coefficient s for such a mutant follows a Γ distribution with the mean $\bar{s}$ and the parameter γ;

$$f_a(s) = \alpha^\gamma e^{-\alpha s} s^{\gamma-1}/\Gamma(\gamma), \qquad [25]$$

in which $\alpha = \gamma/\bar{s}$ and $\gamma > 0$.

Noting that the probability of ultimate fixation of a single mutant with selective advantage s (>0) is $u = 2s(N_e/N)/(1 - e^{-4N_e s})$ (see ref. 11, p. 426), the rate of evolution due to advantageous mutations is

$$k_g = \int_0^\infty 2Nv_a u f_a(s) ds$$

$$= 4N_e v_a \bar{s} R^{\gamma+1} \sum_{j=0}^{\infty} (j + R)^{-\gamma-1}, \qquad [26]$$

in which $R = \gamma/(4N_e\bar{s})$. This can be approximated by

$$k_g = v_a\{4N_e\bar{s} + [\gamma/(6N_e\bar{s} + \gamma)]^\gamma\}. \qquad [27]$$

For $N_e\bar{s} \gg 1$, we have $k \approx 4N_e\bar{s}v_a$. This means that the rate of evolution can become enormously high in a very large population, k_g being directly proportional to N_e, contrary to actual observations.

DISCUSSION

The distribution of selection coefficients of new mutations at the neighborhood of strict neutrality was discussed by Crow (6) and King (7). However, it was Ohta (8) who investigated quantitatively the problem of "near neutrality" by assuming a specific mathematical form of the distribution. In Ohta's model, mutations are assumed to be deleterious and the selection coefficients against individual mutants follow an exponential distribution. With this model, she showed that the level of heterozygosity reaches an upper limit as N_e increases, whereas the rate of evolution per generation (k_g) in terms of mutant substitutions is inversely proportional to the effective population size N_e. The present model assuming a Γ distribution of selective disadvantages of mutants among different sites within a gene (or among different amino acid sites within a protein) has an advantage over Ohta's model in that it can accommodate a much larger fraction of effectively neutral mutations. Actually, Ohta's model corresponds to $\beta = 1$, while the assumption that all the mutations are neutral corresponds to $\beta \to 0$. It is likely that an intermediate parameter value such as $\beta = 0.5$ (as depicted in Fig. 1) may be more realistic to describe the typical situation observed in natural populations. In the case of $\beta = 0.5$, the rate of evolution per generation is inversely proportional to $\sqrt{N_e}$ if $N_e\bar{s}' \gg 1$; i.e., $k_g \propto 1/\sqrt{N_e}$. Thus the evolutionary rate per year is $k_1 \propto 1/(g\sqrt{N_e})$, in which g is the generation span in years. If g is inversely proportional to $\sqrt{N_e}$ among various organisms, then $g\sqrt{N_e}$ is constant, and therefore the evolutionary rate per year is constant, provided that the mutation rate v per generation is constant (uniform) among them.

Note that in the present model those mutations that become fixed in the population by random drift in the course of evolution are restricted to effectively neutral mutations. The selective disadvantage of such mutants is at most of the order of $1/(2N_e)$, which means 10^{-5} or less in many mammals. The proportion of effectively neutral mutants decreases as the population size increases. This is why the heterozygosity increases much more slowly in the present model as compared with the conventional model of neutral mutations (see Fig. 3). The observations that the average heterozygosity is restricted in most organisms to the range 0% to 20% have been used repeatedly as evidence against the neutral mutation theory (see ref. 13). It is likely that this difficulty is resolved by the present model if we assume in addition that a population bottleneck occurs from time to time in all organisms in the course of evolution, reducing their effective population sizes substantially (14). Recently, Li (15, 16) investigated the amount of genetic variability maintained in a finite population using the K allele model incorporating two or three classes of mutations including the neutral and slightly deleterious classes. Similarly, Maruyama and Kimura (17) used a stepwise mutation model incorporating two types of mutations, neutral and slightly deleterious, to investigate the same problem. The present model has a more desirable feature of incorporating a continuous spectrum of mutations conferring different fitnesses, but it does not take into account the limited detection ability of electrophoretic methods. At any rate, the present model can explain the observation made by Ohta (18) that in both *Drosophila* and humans the proportion of rare alleles is greater than what is expected under the assumption that all the mutations are strictly neutral. This observation, if valid, will greatly reduce the utility of Ewens' sampling theory (19), a point that was recently elaborated by Li (20). In the examples illustrated in Figs. 2 and 3, we assume the mutation rate 2×10^{-6} per locus per generation. This value is based on the results reported by Mukai and Cockerham (21) for *Drosophila melanogaster* and by Nei (22) for humans and the Japanese macaque. Recently, higher estimates of mutation rates have been reported by Neel and Rothman (23) for tribal Ameridians.

As to the rate of molecular evolution, the present model with $v = 2 \times 10^{-6}$, $\beta = 0.5$, and $\bar{s}' = 0.001$, as illustrated in Fig. 2, seems to give realistic values; in mammals, $N_e = 10^5$ may be

a representative effective population size for many species during evolution, and $k \approx 10^{-7}$, as shown in Fig. 2, is not very far from the typical rate, which is of the order 1.5×10^{-7} per cistron per year (as represented by globins).

It is likely that the value of the parameter β is smaller in mammals than in insects, because of higher physiological homeostasis in the mammals. The possibility of more mutations being neutral in higher forms such as mammals with advanced homeostasis has been suggested by Kondo (24). Low physiological homeostasis and frequent local extinction of colonies must be the main reason why the heterozygosity (or 1, minus the sum of squares of allelic frequencies in haploid organisms) does not go very high in organisms having immense apparent population sizes such as neotropical *Drosophila* (25) and *Escherichia coli* (26). The mathematical model proposed in this paper represents my attempt to make the neutral mutation theory more precise and realistic. The model assumes that molecular evolution and polymorphism are caused by random drift of very slightly deleterious but effectively neutral mutations. In this respect, the present theory resembles Ohta's theory of slightly deleterious mutations (27–29). But there are some important differences. Ohta (29) claims that, in very large populations, the stable mutation–selection balance will be realized with heterozygosity reaching the upper limit, while molecular evolution should have stopped or at least have slowed down. Then, fixation of mutants is mainly restricted to population bottlenecks at the time of speciation. On the other hand, I assume that, even in very large populations, alleles at intermediate frequencies, as often found in *Drosophila* species (see ref. 30, table II), represent effectively neutral mutations carried by random drift and that evolution by drift is unlikely to be stopped in these species. Finally, there is one biological problem that we have to consider. Under the present model, effectively neutral, but, in fact, very slightly deleterious mutants accumulate continuously in every species. The selective disadvantage of such mutants (in terms of an individual's survival and reproduction—i.e., in Darwinian fitness) is likely to be of the order of 10^{-5} or less, but with 10^4 loci per genome coding for various proteins and each accumulating the mutants at the rate of 10^{-6} per generation, the rate of loss of fitness per generation may amount to 10^{-7} per generation. Whether such a small rate of deterioration in fitness constitutes a threat to the survival and welfare of the species (not to the individual) is a moot point, but this will easily be taken care of by adaptive gene substitutions that must occur from time to time (say once every few hundred generations).

I thank Dr. Tomoko Ohta and Professor Masao Kotani for stimulating discussions. Thanks are also due to Drs. J. F. Crow, Alan Robertson, and Roger Milkman, who read the first draft and made many useful suggestions. This is contribution no. 1252 from the National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan.

1. Kimura, M. (1969) *Proc. Natl. Acad. Sci. USA* **63,** 1181–1188.
2. Kimura, M. & Ohta, T. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 2848–2852.
3. Fuerst, P. A., Chakraborty, R. & Nei, M. (1977) *Genetics* **86,** 455–483.
4. Kimura, M. (1968) *Nature (London)* **217,** 624–626.
5. Kimura, M. & Ohta, T. (1971) *Nature (London)* **229,** 467–469.
6. Crow, J. F. (1972) in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability,* eds. LeCam, L. M., Neyman, J. & Scott, E. L. (Univ. of California Press, Berkeley), Vol. 5, pp. 1–22.
7. King, J. L. (1972) in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability,* eds. LeCam, L. M., Neyman, J. & Scott, E. L. (Univ. of California Press, Berkeley), Vol. 5, pp. 69–100.
8. Ohta, T. (1977) in *Molecular Evolution and Polymorphism,* ed. Kimura, M. (Natl. Inst. of Genet., Mishima, Japan), pp. 148–167.
9. Fisher, R. A. (1930) *The Genetical Theory of Natural Selection* (Clarendon, Oxford).
10. Kimura, M. (1969) *Genetics* **61,** 893–903.
11. Crow, J. F. & Kimura, M. (1970) *An Introduction to Population Genetics Theory* (Harper & Row, New York).
12. Ohta, T. & Kimura, M. (1973) *Genet. Res.* **22,** 201–204.
13. Lewontin, R. C. (1974) *The Genetic Basis of Evolutionary Change* (Columbia Univ. Press, New York).
14. Nei, M., Maruyama, T. & Chakraborty, R. (1975) *Evolution* **29,** 1–10.
15. Li, W.-H. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 2509–2513.
16. Li, W.-H. (1978) *Genetics* **90,** 349–382.
17. Maruyama, T. & Kimura, M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 919–922.
18. Ohta, T. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3194–3196.
19. Ewens, W. J. (1972) *Theor. Popul. Biol.* **3,** 87–112.
20. Li, W.-H. (1979) *Genetics,* in press.
21. Mukai, T. & Cockerham, C. C. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 2514–2517.
22. Nei, M. (1977) *Am. J. Hum. Genet.* **29,** 225–232.
23. Neel, J. V. & Rothman, E. D. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5585–5588.
24. Kondo, S. (1977) in *Molecular Evolution and Polymorphism,* ed. Kimura, M. (Natl. Inst. of Genet., Mishima, Japan), pp. 313–331.
25. Ayala, F. J., Powell, J. R., Tracey, M. L., Mourão, C. A. & Pérez-Salas, S. (1972) *Genetics* **70,** 113–139.
26. Milkman, R. (1973) *Science* **182,** 1024–1026.
27. Ohta, T. (1973) *Nature (London)* **246,** 96–98.
28. Ohta, T. (1974) *Nature (London)* **252,** 351–354.
29. Ohta, T. (1976) *Theor. Popul. Biol.* **10,** 254–275.
30. Selander, R. K. (1976) in *Molecular Evolution,* ed. Ayala, F. J. (Sinauer, Sunderland, MA), pp. 21–45.