Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5773–5777, September 1981
Genetics

Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons

(molecular evolution/quantitative character/natural selection/neutral mutation–random drift hypothesis/mode of evolution)

MOTOO KIMURA

National Institute of Genetics, Mishima, 411 Japan

Contributed by Motoo Kimura, June 3, 1981

ABSTRACT **The rate of evolution in terms of the number of mutant substitutions in a finite population is investigated assuming a quantitative character subject to stabilizing selection, which is known to be the most prevalent type of natural selection. It is shown that, if a large number of segregating loci (or sites) are involved, the average selection coefficient per mutant under stabilizing selection may be exceedingly small. These mutants are very slightly deleterious but nearly neutral, so that mutant substitutions are mainly controlled by random drift, although the rate of evolution may be lower as compared with the situation in which all the mutations are strictly neutral. This is treated quantitatively by using the diffusion equation method in population genetics. A model of random drift under stabilizing selection is then applied to the problem of "nonrandom" or unequal usage of synonymous codons, and it is shown that such nonrandomness can readily be understood within the framework of the neutral mutation–random drift hypothesis (the neutral theory, for short) of molecular evolution.**

It is generally accepted that stabilizing selection is the most prevalent type of natural selection at the phenotypic level (1–4). It eliminates phenotypically extreme individuals and preserves those that are near the population mean (5). It is also called centripetal selection (6) or normalizing selection (7), and many examples have been reported. Probably the best example in human populations is the relationship between the birth weights of babies and their neonatal mortality, as studied by Karn and Penrose (8). These authors found that babies whose weight is very near the mean have the lowest mortality. This optimum weight is slightly heavier than the mean, and mortality increases progressively as the birth weight deviates from this optimum (see also ref. 9). Unlike the type of natural selection that Darwin (10) had in mind when he tried to explain evolution, stabilizing selection acts to keep the status quo rather than to cause a directional change. From this, it might appear that stabilizing selection is antithetical to evolutionary change.

In this note, I intend to show that, under stabilizing phenotypic selection, extensive "neutral evolution" can occur. By neutral evolution, I mean accumulation of mutant genes in the species through random genetic drift (due to finite population size) under mutational pressure. Thus, beneath an unchanged morphology, a great deal of cryptic genetic change may be occurring in natural populations of all organisms, transforming even genes of "living fossils" (11). This will substantiate my neutral mutation–random drift hypothesis (the neutral theory, for short) of molecular evolution (ref. 12; for review, see ref. 13). I shall also show that this gives a clue to understanding "nonrandom" or unequal usage of synonymous codons (14–16) based on the neutral theory.

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.

Selection intensity at an individual locus when overall phenotypic selection is given

Let us consider a quantitative character, such as height, weight, concentration of some substance, or a more abstract quantity that represents Darwinian fitness in an important way. We assume that the character is determined by a large number of loci (or sites), each with a very small effect in addition to being subjected to environmental effects. We also assume that genes are additive with respect to the character. We follow the method used by Bulmer (17, 18) and Kimura and Crow (19). Let X be the measured phenotypic character with the mean M and the variance σ^2. We denote by X_{op} the optimum phenotypic value and, unless otherwise stated, we shall take this point as the origin. Let $F(X)$ and $W(X)$ be the relative frequency and the fitness of individuals with character value X. Two examples of $W(X)$ are shown in Fig. 1.

Consider a particular locus at which a pair of alleles A_1 and A_2 are segregating with respective frequencies $1 - p$ and p. We assume a random-mating diploid population and let X_{ij} be the average phenotypic value of A_iA_j individuals, where $i = 1$ or 2 and $j = 1$ or 2.

It is often convenient to measure various quantities relating to the character value in units of the standard deviation (σ). For this purpose, lowercase letters will be used such as $x = (X - X_{op})/\sigma$, and the corresponding frequency and fitness functions will be denoted by $f(x)$ and $w(x)$. We also let $m = (M - X_{op})/\sigma$ and $a_{ij} = (X_{ij} - M)/\sigma$. Note that a_{ij} is the deviation of A_iA_j from the population mean in σ units.

We assume that the background distribution of the character is the same among different genotypes at this locus and that this is given by $f(x)$ with good approximation, because individual gene effects are assumed to be extremely small. Let w_{ij} be the relative fitness of A_iA_j, then

$$w_{ij} = \int_{-\infty}^{\infty} w(x)f(x - a_{ij})dx, \qquad [1]$$

as explained in ref. 19. Assuming that a_{ij} is small, we expand $f(x - a_{ij})$ in a Taylor series,

$$f(x - a_{ij}) = f(x) - a_{ij}f'(x) + (a_{ij}^2/2)f''(x) - \ldots, \qquad [2]$$

as in ref. 19, but here we retain the second-order term, so that we get, from Eq. 1,

$$w_{ij} = b_0 - a_{ij}b_1 + a_{ij}^2 b_2/2, \qquad [3]$$

where

$$b_0 = \int_{-\infty}^{\infty} w(x)f(x)dx, \qquad b_1 = \int_{-\infty}^{\infty} w(x)f'(x)dx,$$

and

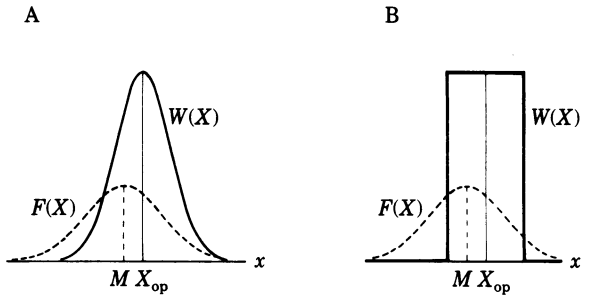

FIG. 1. Examples of the fitness function $W(X)$. (*A*) Normal distribution. (*B*) Uniform distribution. ——, fitness function; - - - -, frequency function.

$$b_2 = \int_{-\infty}^{\infty} w(x)f''(x)dx.$$

Here, the prime denotes differentiation.

Let a be the effect of substituting A_2 for A_1 on the character value x. Then, under random mating and assuming an additive gene effect on x, we find that $a_{11} = -2ap$, $a_{12} = a(1 - 2p)$, and $a_{22} = 2a(1 - p)$. Then, by using Eq. **3**, we can compute the mean fitness $\bar{w} = w_{11}(1 - p)^2 + w_{12}2(1 - p)p + w_{22}p^2$ and the average fitness of A_2—i.e., $w_2 = w_{21}(1 - p) + w_{22}p$—and they turn out to be as follows:

$$\bar{w} = b_0 + a^2p(1 - p)b_2 \quad [4]$$

$$w_2 = b_0 - a(1 - p)b_1 + a^2(1 - p)b_2/2. \quad [5]$$

The change of the frequency of A_2 in one generation is given by $\Delta p = p(w_2 - \bar{w})/\bar{w}$ (see p. 180 of ref. 20). Therefore, by substituting Eqs. **4** and **5** in this expression and neglecting terms involving a^3 and higher order terms, we obtain

$$\Delta p = \frac{p(1 - p)}{b_0}\left[-ab_1 + a^2b_2\left(\frac{1}{2} - p\right)\right], \quad [6]$$

in agreement with Bulmer (17). Then, the selection coefficient, s, which represents the selective advantage of A_2 over A_1 is

$$s = -ab_1/b_0 + a^2b_2(\tfrac{1}{2} - p)/b_0. \quad [7]$$

With this selection coefficient s, the change of p by selection per generation is $\Delta p = sp(1 - p)$.

In the special, but important, case in which both the frequency and the fitness functions are given by normal distributions,

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp[-(x - m)^2/2], \quad [8]$$

and

$$w(x) = \exp(-kx^2), \quad [9]$$

Eq. **7** reduces to

$$s = -m\lambda a + (\lambda^2m^2 - \lambda)(\tfrac{1}{2} - p)a^2, \quad [10]$$

where $\lambda = 2k/(1 + 2k)$. An equivalent result was obtained earlier by Bulmer (18).

Note that, if we use the original scale (X) and express the density function of the frequency distribution of the character by $F(X) = (1/\sqrt{2\pi\sigma})\exp[-(X - M)^2/2\sigma^2]$, the fitness function by $W(X) = \exp(-KX^2)$, and the effect of allele substitution by A, then the parameters in Eq. **10** are $m = (M - X_{op})/\sigma$, $\lambda = 2K\sigma^2/(1 + 2K\sigma^2)$, and $a = A/\sigma$. From Eq. **10**, it may be seen that, if the deviation of the mean from the optimum is much larger than the effect of the allele substitution ($|m| \gg |a|$), we have $s/a \approx -\lambda m$, which agrees with equation **23** of ref. 19. In this case, the situation is similar to truncation selection (see refs. 21 and 22), and natural selection acts very efficiently to change the mean toward the optimum. During this short period of directional selection, extensive shift of gene frequencies is expected to occur at many loci, but this process itself will seldom cause gene substitutions.

If, on the other hand, the mean is at the optimum ($m = 0$), we have $s = \lambda(\tfrac{1}{2} - p)a^2$ from Eq. **10**. In this case, A_2 is disadvantageous if $p < \tfrac{1}{2}$ but advantageous if $p > \tfrac{1}{2}$. This selection is frequency dependent, and alleles behave as if negatively overdominant. The change of gene frequency is then given by $\Delta p = \lambda a^2p(1 - p)(p - \tfrac{1}{2})$ in agreement with Robertson (23) and Wright (24). What is pertinent to our evolutionary consideration is that, here, every mutation is deleterious, because $\Delta p < 0$ if p is small. Furthermore, if a large number of loci are segregating, each with a very small effect, a is small so that a^2 is extremely small. This applies with particular force if we consider an individual nucleotide site rather than the conventional gene locus, as it has been estimated (25) that the average individual in a large mammalian population is likely to be heterozygous at a million or so nucleotide sites. This substantiates Ohta's claim (26, 27) that the majority of mutants at the molecular level are nearly neutral but very slightly deleterious. As I shall show below, negatively overdominant alleles are far more susceptible to random genetic drift than unconditionally deleterious alleles having the same magnitude of selection coefficient.

As to the intensity of natural selection involved, we can measure it in terms of load (L)—i.e., by the fraction of individuals that are eliminated in each generation by natural selection due to deviation of their phenotypic values from the optimum. For the frequency and fitness functions given as Eqs. **8** and **9**, we obtain

$$L = 1 - \sqrt{1 - \lambda}\exp(-\lambda m^2/2). \quad [11]$$

For $m = 0$, this reduces to $L = 1 - \sqrt{1 - \lambda}$. In the special case in which the fitness function has the same variance as the frequency function—i.e., when $K = 1/(2\sigma^2)$ or $k = \tfrac{1}{2}$, we get $L \approx 0.293$ or $\approx 30\%$ elimination. In general, L is likely to be small for any single character in mammals. For example, according to Haldane (1), the intensity of selection acting on birth weight of babies through their neonatal mortality is $L = 0.027$. If L is small, we have, with good approximation, $L = \lambda/2$.

Behavior of mutant alleles in a finite population under stabilizing selection

Probability of Fixation of a Mutant Allele. We denote by N_e the effective size of the population (for the meaning of N_e, see ref. 28). Roughly speaking, N_e is equal to the number of breeding individuals in one generation. This number is likely to be much smaller in most cases than the apparent population size, which we denote by N. To simplify expressions, we let $\beta_1 = -\lambda ma$ and $\beta_2 = \lambda(1 - \lambda m^2)a^2/2$, so that Eq. **10** reduces to

$$s = \beta_1 - \beta_2(1 - 2p), \quad [12]$$

where p is the frequency of A_2. Ignoring mutational change for a moment, and denoting the frequency of A_2 by y, the mean and the variance in the change of y during one generation are

$$M_{\delta y} = [\beta_1 - \beta_2(1 - 2y)]\,y(1 - y) \quad [13]$$

and

$$V_{\delta y} = y(1 - y)/2N_e. \quad [14]$$

Let $u(p)$ be the probability that A_2 becomes eventually fixed in the population (i.e., reaches 100% in frequency), given that its initial frequency is p. Then, $u(p)$ can be expressed in terms of $M_{\delta y}$ and $V_{\delta y}$ by using a general formula for the probability of fixation obtained by Kimura (ref. 29; see also p. 424 of ref. 20, where x is used instead of y). We are particularly interested in the probability of fixation of A_2 when it is initially singly represented in the population. If we denote this probability by u, then this is given by $u(p)$ with $p = 1/(2N)$. We then obtain

$$u = 1 \Big/ \left[2N \int_0^1 \exp\{-B_1 x + B_2 x(1-x)\} dx \right], \qquad [15]$$

where $B_1 = 4N_e\beta_1$ and $B_2 = 4N_e\beta_2$.

In the above treatment, we have assumed that m (the deviation of the mean from the optimum) remains unchanged throughout the process. This assumption appears to be unrealistic because, if $m \neq 0$, one would expect $|m|$ to be reduced with time by the directional component of selection. There is an important possibility, however, that this change is opposed by mutational pressure so that m remains constant under continued stabilizing selection, although $|m|$ at equilibrium is likely to be small. This occurs when the optimum and the mutational equilibrium point do not coincide. We shall elaborate such a case when we discuss the problem of nonrandom synonymous codon usage.

To show that mutants that have negative overdominance (as induced by stabilizing selection) are far more likely to be fixed by random drift than unconditionally deleterious mutants that have comparable selection coefficients, some examples of the probabilities of fixation (u) for these two cases are listed in Table 1. In the case of stabilizing selection, we let $m = 0$ and denote $\lambda a^2/2$ by s_s (selection coefficient for stabilizing selection), so that $\Delta p = -s_s p(1-p)(1-2p)$. For the unconditionally deleterious case, we denote the selection coefficient against A_2 by $-s'$ ($s' > 0$), so that the probability of fixation is given by

$$u = S'/[2N(e^{S'} - 1)], \qquad [16]$$

where $S' = 4N_e s'$ (see p. 426 of ref. 20). In both cases, u is tabulated taking the probability of fixation of the completely neutral case as the unit—i.e., it is expressed as a multiple of $u_0 = 1/(2N)$. It is clear from Table 1 that an enormous difference exists between the two cases in fixation probability and that, under stabilizing selection, extensive neutral evolution is possible even when $4N_e s_s$ is 8 or more. For $B = 4N_e s_s > 8$, it can be shown that $u/u_0 \approx \sqrt{B/\pi}\exp(-B/4)$.

Gene Frequency Distribution. We now incorporate mutational pressure and investigate the probability distribution of allelic frequencies at statistical equilibrium attained under stabilizing selection in a finite population. We shall denote by $\phi(p)$ the probability density such that $\phi(p)dp$ represents the probability that the frequency of A_2 in the population lies in the range $p \sim p + dp$, where $0 < p < 1$. We assume reversible mutations between the two alleles and let v_1 be the mutation rate from A_1 to A_2 and v_2 be the rate in the reverse direction. Then, the mean and the variance in the change of A_2 in one generation are, respectively,

$$M_{\delta p} = [\beta_1 - \beta_2(1-2p)]\,p(1-p) - v_2 p + v_1(1-p) \qquad [17]$$

and

$$V_{\delta p} = p(1-p)/(2N_e). \qquad [18]$$

By using Wright's (30) formula for the steady-state gene-frequency distribution (see p. 434 of ref. 20), we obtain

$$\phi(p) = Ce^{B_1 p - B_2 p(1-p)}\, p^{V_1 - 1}(1-p)^{V_2 - 1}, \qquad [19]$$

where $B_1 = 4N_e\beta_1$, $B_2 = 4N_e\beta_2$, $V_1 = 4N_e v_1$, $V_2 = 4N_e v_2$, and C is determined so that $\int_0^1 \phi(p)dp = 1$. The probability of A_2 being temporarily fixed in the population may be obtained by integrating $\phi(p)$ from $1 - [1/(2N)]$ to 1, and we obtain, with sufficient accuracy,

$$f_2 = C\,e^{B_1}/[V_2(2N)^{V_2}]. \qquad [20]$$

Similarly, the probability of A_1 being temporarily fixed in the population (i.e., A_2 lost) is

$$f_1 = C/[V_1(2N)^{V_1}]. \qquad [21]$$

Then, the ratio of f_2 to f_1 is

$$f_2/f_1 = e^{B_1}(V_1/V_2)(2N)^{V_1 - V_2}. \qquad [22]$$

In this paper, we shall be mainly concerned with the situation in which both V_1 and V_2 are much smaller than unity and alleles are fixed most of the time. This situation is particularly pertinent when we consider individual nucleotide sites rather than conventional gene loci, because the mutation rate per site must be of the order of 10^{-8} rather than of 10^{-5}.

In general, for any set of values of B_1, B_2, V_1, and V_2, we can compute the mean frequency $\bar{p}$ and the mean heterozygosity $\bar{H}_e$ per locus through numerical integration by using $\bar{p} = E(p) = \int_0^1 p\phi(p)dp$ and $\bar{H}_e = E[2p(1-p)] = \int_0^1 2p(1-p)\phi(p)dp$. If the phenotype is determined by n equivalent loci in addition to environmental effects, we have $M = 2nA\bar{p}$ and $\sigma^2 = nE[2A^2p(1-p)]/\rho^2 = nA^2\bar{H}_e/\rho^2$, where A is the effect of substituting A_2 for A_1 on the character ($A = a\sigma$) and ρ^2 is the fraction of phenotypic variance due to gene segregation—i.e., broad sense heritability. Furthermore, if mutation rates are equal in both directions ($v_1 = v_2 \equiv v$) and the phenotypic mean coincides with the optimum phenotype (i.e., $m = 0$ or $B_1 = 0$, $B_2 = 2N_e\lambda a^2$), the distribution formula is much simplified, and the values of $E[p(1-p)]$ are tabulated by Bulmer (18) for some combinations of values of $4N_e v$ and $v/(ka^2)$. One interesting property of the frequency distribution of alleles under stabilizing selection is that it is more U shaped than the strictly neutral case having the same mutation rate. It may sometimes be convenient to take as the standard the situation in which the minus alleles are fixed at all loci. Then, the range of X lies between 0 and $2nA$ if we assume that n loci are involved and the effect of allele substitution is the same at all the loci. Let

$$X_{op} = 2nAQ_{op}, \qquad [23]$$

where Q_{op} is the position of the optimum when the total range of X is rescaled so that it lies in the interval [0, 1]. If A_1 is the minus allele and A_2 is the plus allele, so that $A > 0$, then the optimum is less than the mean if $Q_{op} < \bar{p}$ and more than the mean if $Q_{op} > \bar{p}$.

Table 1. Relative probability of fixation (u/u_0) of negatively overdominant and unconditionally deleterious mutants

$\tilde{S}$	Negatively overdominant	Unconditionally deleterious
0	1.00	1.00
1.0	0.84	0.58
8.0	0.23	2.7×10^{-3}
16.0	0.042	1.8×10^{-6}
20.0	0.017	4.1×10^{-8}
30.0	0.0017	2.8×10^{-12}

$\tilde{S}$ stands for $4N_e s_s$ for the negatively overdominant case and $4N_e s'$ for the unconditionally deleterious case.

Application to the problem of nonrandom codon usage

Recently, nonrandom or unequal usage of synonymous codons has been reported in many genes of various organisms (for review, see ref. 14). Indeed, nonrandom codon usage appears to be a rule rather than an exception, and this is often mentioned as evidence against the neutral theory. I shall now show that this can be explained in the framework of the neutral theory. Note that the existence of selective constraint (negative selection) by no means contradicts the neutral theory (see ref. 13).

To simplify the argument, we group nucleotide bases A (adenine) and U (uracil) as A_1 and C (cytosine) and G (guanine) as A_2. It is known (31) that, at the third position of degenerate codons in mammalian mRNAs, A_2 predominate over A_1. For globin mRNA, the ratio of A_2 to A_1 at position 3 is ≈7:3 (32). As shown above, the distribution function $\phi(p)$ (Eq. 19), when applied to a nucleotide site rather than a gene locus, indicates that either A_1 or A_2 is fixed most of the time in the course of evolution. This is because the mutation rate per site is exceedingly low, so that the probability of polymorphism per nucleotide site is very small, although this probability may amount to more than 10% when applied to a locus that is comprised of 1000 or so nucleotide sites.

As to the cause of nonrandom codon usage, recent studies of Ikemura (15, 16) are instructive. He found a strong positive correlation between the frequency of synonymous codon usage and abundance of cognate tRNA in *Escherichia coli*. This correlation appears to be related to the translational efficiency (see also ref. 33). If this applies in general to other organisms, the most plausible explanation for preferential codon usage is that it represents the optimum state in which the population of synonymous codons matches that of cognate tRNA available in the cell. This will help to carry out more efficient cell function, leading to higher Darwinian fitness. This appears to be compatible with the genome hypothesis of Grantham *et al.* (14), who claim that a surprising consistency of choices of degenerate bases exists among genes of the same or similar genomes and that "the genome and not the individual gene is the unit of selection."

Let Q_{op} be the optimum proportion of A_2 (guanine or cytosine) at position 3 of the codons and assume that mutation rates are equal between A_1 and A_2—i.e., $V_1 = V_2$—then the mean of $p(\bar{p})$ does not coincide with Q_{op} unless $Q_{op} = 0.5$. So, we assume that stabilizing selection is at work to hold $\bar{p}$ near Q_{op}. At individual sites, however, A_2 is either fixed or lost most of the time. Let f_2 be the probability that A_2 is fixed in the population at a given site. Similarly, let f_1 be the probability that A_1 is fixed (A_2 is lost). Then, from Eq. 22, we have $f_2/f_1 = \exp(B_1)$, where $B_1 = 4N_e\beta_1 = -4N_e\lambda ma$. Thus, we can estimate B_1 by the relationship $B_1 = \ln(f_2/f_1)$, and we obtain $B_1 = 0.85$ for $f_2/f_1 = 0.7/0.3$. In most mammalian species, the effective size N_e must be at least 10^4. Therefore, the intensity of selection that acts at an individual site to produce nonrandom codon usage is an exceedingly weak one, leaving plenty of room for random drift to operate. This is consistent with Latter's (34) claim that mutations responsible for enzyme polymorphisms are very slightly deleterious with "*Ns*" values in the range 1–3.

One important question that remains is the extent to which the rate of evolution in terms of mutant substitution is influenced by such selection. As the relative evolutionary rate (in terms of mutant substitution) under stabilizing selection as compared with the strictly neutral case is given by u/u_0 with $u_0 = 1/(2N)$, we have, from Eq. 15,

$$u/u_0 = 1 \Big/ \int_0^1 \exp[-B_1x + B_2x(1-x)]dx, \qquad [24]$$

where $B_1 = -4N_e\lambda ma$ and $B_2 = 2N_e\lambda(1 - \lambda m^2)a^2$. If we assume that $2N_e\lambda a^2$ is negligibly small, so that $B_2 \approx 0$, then we get

$$u/u_0 \approx 2f_1f_2\ln(f_2/f_1)/(f_2 - f_1). \qquad [25]$$

For $f_2/f_1 = 0.7/0.3$, as we observe at the third position of the codons in globin and other mammalian mRNAs, we get $u/u_0 = 0.89$. In other words, the evolution is retarded by ≈10% from what is expected under complete selective neutrality. Under the more extreme condition $f_2/f_1 = 0.9/0.1$, we get $u/u_0 \approx 0.49$, which means ≈50% retardation.

In actual situations, however, there are four possible "alleles" (bases) per nucleotide site rather than two and, together with other complications due to differences in the speed of translation among different types of genes, etc. (16), we need more careful and detailed analysis to arrive at a more accurate figure for the retardation.

Discussion

During its lifetime, an individual is subject to natural selection through a large number of quantitative characters, many of which are mutually correlated. Let us assume, to simplify the treatment, that we can choose a certain number, say n_c, of independent characteristics that collectively represent, to a first approximation, the total pattern of selection. Various parameters pertaining to the ith character will be expressed by subscript i ($i = 1, 2, \ldots, n_c$).

Because the total selection intensity is limited, the selection intensity, as measured by L_i at each component character is expected to be small if the number of characters involved is large. Let L_T be the total selection intensity, then $(1 - L_T) = \Pi_i(1 - L_i)$, so that $L_T \approx 1 - \exp(-\Sigma_i L_i)$. To simplify the treatment still further, let us suppose that the L_is are all equal among component characters, so that $L_i \approx -(1/n_c)\ln(1 - L_T)$. The selection coefficient per site is then $-\lambda_i(1/2 - p_i)a_i^2$ and, noting that $L_i \approx \lambda_i/2$ and dropping the subscript i, we have $s = \ln(1 - L_T)(1 - 2p)a^2/n_c$ (approximately). On the right-hand side of this formula, we note that $a^2/n_c = A^2/(n_c\sigma^2)$, where $n_c\sigma^2$ is the variance of the total phenotype. Let $\bar{h}_e$ be the average heterozygosity per site and, if we denote by n_{nuc} the total number of nucleotide sites concerned, then $n_{nuc}A^2\bar{h}_e = n_c\sigma^2\rho^2$, where A is the effect of substituting one nucleotide on a component phenotype and ρ^2 is broad sense heritability. Thus, the coefficient for stabilizing selection s_s, as defined by the relationship $s = -s_s(1 - 2p)$, turns out to be $s_s = -[\ln(1 - L_T)]\rho^2/(n_{nuc}\bar{h}_e)$. This represents the selection intensity involved in nucleotide substitution under stabilizing selection (assuming $m = 0$).

Let us assume that the average heterozygosity per enzyme locus with respect to electrophoretically detectable alleles is 0.1 and (rather conservatively) that there is twice as much heterozygosity with respect to silent alleles. Then, if the average number of nucleotide sites that comprises a locus is 10^3, we get $\bar{h}_e = 3 \times 10^{-4}$. Extrapolating this to the total genome of a mammal that has 3.5×10^9 nucleotide sites, the average number of heterozygous nucleotide sites per individual is $n_{nuc}\bar{h}_e = 1.05 \times 10^6$. As typical values of genetic load and heritability for a mammal (such as the human species), let us suppose that $L_T = 0.5$ and $\rho^2 = 0.5$, then, we obtain $s_s = 3.3 \times 10^{-7}$. This is a very small selection coefficient for stabilizing selection and shows that the majority of mutations at the molecular level are nearly neutral but very slightly deleterious. This agrees with Ohta's hypothesis of very slightly deleterious mutations (26, 27). However, the fitness of the species does not drift downward in this view as it does in Ohta's hypothesis. Also, in this view, those genes that are substituted by random drift and those that are responsible for phenotypic variability of quantitative traits belong to the same class. It is possible that many, and even most, of the mu-

tants affecting a quantitative trait are regulatory rather than structural. DNA outside the coding region may be more important from this standpoint than translated DNA. The present analysis agrees with Lande (35), who suggests that many polygenic changes can accumulate by random drift because they have little or no net phenotypic effect.

Needless to say, some sites produce much larger phenotypic effects than others and therefore are subject to stronger selection. On the other hand, a certain fraction of sites (presumably a large fraction) produce no phenotypic effects at all and therefore are completely neutral with respect to natural selection.

The picture of evolution that emerges from the present analysis is as follows. From time to time, the position of the optimum shifts due to changes in environment and the species tracks such changes rapidly by altering its mean. But, most of the time, stabilizing selection predominates. Under this selection, neutral evolution (random fixation of alleles by sampling drift) occurs extensively, transforming all genes, including those of living fossils, profoundly at the molecular level.

I thank Drs. Tomoko Ohta and R. Milkman for stimulating discussions in composing the manuscript. Thanks are also due to Drs. J. Maynard Smith, J. F. Crow, and K. Aoki for helpful suggestions and useful comments to improve the presentation. This is contribution no. 1364 from the National Institute of Genetics, Mishima, Shizuoka-ken, 411 Japan.

1. Haldane, J. B. S. (1959) in *Darwin's Biological Work*, ed. Bell, P. R. (Cambridge Univ. Press, Cambridge, England), pp. 101–149.
2. Mather, K. (1973) *Genetical Structure of Populations* (Chapman & Hall, London).
3. Wright, S. (1977) *Evolution and the Genetics of Populations* (Univ. Chicago Press, Chicago), Vol. 3.
4. Parkin, D. T. (1979) *An Introduction to Evolutionary Genetics* (Arnold, London).
5. Mather, K. (1953) *Symp. Soc. Exp. Biol.* **7**, 66–95.
6. Simpson, G. G. (1944) *Tempo and Mode in Evolution* (Columbia Univ. Press, New York).
7. Waddington, C. H. (1957) *The Strategy of the Genes* (Allen & Unwin, London).
8. Karn, M. N. & Penrose, L. S. (1951) *Ann. Eugen.* **16**, 147–164.
9. Cavalli-Sforza, L. L. & Bodmer, W. F. (1971) *The Genetics of Human Populations* (Freeman, San Francisco).
10. Darwin, C. (1859) *The Origin of Species* (Murray, London).
11. Kimura, M. (1969) *Proc. Natl. Acad. Sci. USA* **63**, 1181–1188.
12. Kimura, M. (1968) *Nature (London)* **217**, 624–626.
13. Kimura, M. (1979) *Sci. Am.* **241** (5), 94–104.
14. Grantham, R., Gautier, C., Gouy, M., Mercier, R. & Pavé, A. (1980) *Nucleic Acids Res.* **8**, r49–r62.
15. Ikemura, T. (1980) in *Genetics and Evolution of RNA Polymerase, tRNA and Ribosomes*, ed. Osawa, S., Ozeki, H., Uchida, H. & Yura, T. (Univ. Tokyo Press, Tokyo), pp. 519–523.
16. Ikemura, T. (1981) *J. Mol. Biol.* **146**, 1–21.
17. Bulmer, M. G. (1971) *Heredity* **27**, 157–162.
18. Bulmer, M. G. (1972) *Genet. Res.* **19**, 17–25.
19. Kimura, M. & Crow, J. F. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 6168–6171.
20. Crow, J. F. & Kimura, M. (1970) *An Introduction to Population Genetics Theory* (Harper & Row, New York); reprinted (1977) by Burgess, Minneapolis, MN.
21. Milkman, R. (1978) *Genetics* **88**, 391–403.
22. Crow, J. F. & Kimura, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 396–399.
23. Robertson, A. (1956) *J. Genetics* **54**, 236–248.
24. Wright, S. (1935) *J. Genetics* **30**, 243–256.
25. Kimura, M. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 515–524.
26. Ohta, T. (1973) *Nature (London)* **246**, 96–98.
27. Ohta, T. (1974) *Nature (London)* **252**, 351–354.
28. Kimura, M. & Ohta, T. (1971) *Theoretical Aspects of Population Genetics* (Princeton Univ. Press, Princeton, NJ).
29. Kimura, M. (1962) *Genetics* **47**, 713–719.
30. Wright, S. (1938) *Proc. Natl. Acad. Sci. USA* **24**, 253–259.
31. Jukes, T. H. (1978) *J. Mol. Evol.* **11**, 121–127.
32. Kimura, M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 454–458.
33. Post, L. E., Strycharz, G. D., Nomura, M., Lewis, H. & Dennis, P. P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1697–1701.
34. Latter, B. D. H. (1975) *Nature (London)* **257**, 590–592.
35. Lande, R. (1980) *Am. Nat.* **116**, 463–479.