Theoretical Study of Near Neutrality. I. Heterozygosity and Rate of Mutant Substitution

Tomoko Ohta and Hidenori Tachida

National Institute of Genetics, Mishima 411, Japan

Manuscript received January 18, 1990
Accepted for publication May 10, 1990

ABSTRACT

In order to clarify the nature of "near neutrality" in molecular evolution and polymorphism, extensive simulation studies were performed. Selection coefficients of new mutations are assumed to be small so that both random genetic drift and selection contribute to determining the behavior of mutants. The model also incorporates normally distributed spatial fluctuation of selection coefficients. If the system starts from "average neutrality," it will move to a better adapted state, and most new mutations will become "slightly deleterious." Monte Carlo simulations have indicated that such adaptation is attained, but that the rate of such "progress" is very low for weak selection. In general, the larger the population size, the more effective the selection becomes. Also, as selection becomes weaker, the behavior of the mutants approaches that of completely neutral genes. Thus, the weaker the selection, the smaller is the effect of population size on mutant dynamics. Increase of heterozygosity with population size is very pronounced for subdivided populations. The significance of these results is discussed in relation to various observed facts on molecular evolution and polymorphism, such as generation-time dependency and overdispersion of the molecular clock, or contrasting patterns of DNA and protein polymorphism among some closely related species.

THE neutral theory of molecular evolution [see KIMURA (1983) for review] now provides a powerful framework for a theoretical interpretation of the evolution of DNA sequences. However, there still remain several unsolved problems relating to the neutral theory. As has been recognized for many years, the neutral theory has two aspects, molecular evolution and polymorphism, and the problems are related to both phases. If the neutral theory is valid, the generation time effect (LI, LUO and WU 1985; SHARP 1989) and the large variance of evolutionary rate among species (GILLESPIE 1986, 1987) should faithfully reflect the mutation rate per year. The other problem is whether the pattern of within-population DNA polymorphism, for which data are rapidly accumulating, fits the neutral theory (KREITMAN 1987; LANGLEY and AQUADRO 1987; AQUADRO, LADO and NOON 1988).

All these problems are much related to the nature of "nearly neutral" mutations, since even mild selection has significant effects on the rate of substitution and polymorphism. Behavior of such mutants is influenced by both selection and drift, but the interaction of the two processes is not yet fully understood. Under the strictly neutral theory, it is expected that evolutionary rate is equal to the neutral mutation rate, and that heterozygosity increases with effective population size (KIMURA 1983). Available data on protein evolution suggest that evolutionary rates are similar between organisms whose generation lengths are greatly different, and that there is an upper limit to virtual heterozygosity of a species (KIMURA 1983; LI, LUO and WU 1985; NEI 1987). Do these facts imply that the mutation rate per year is so similar among organisms, and that the effective population size never becomes so big as to give high virtual heterozygosity beyond the limit? We proposed that, if a class of mutations that are nearly neutral but very slightly deleterious, substantially contributes to molecular evolution and polymorphism, the above curious facts may be satisfactorily understood (OHTA 1972a,b; 1973; 1974; 1977; 1987). This theory is based on the following consideration. Organisms with short generation times tend to be small and to maintain large population sizes. This relationship is thought to hold roughly when long-term effective size is taken into account. Then the proportion of effectively neutral mutations, and hence the evolutionary rate, is decreased, since selection becomes more effective in large populations. If the mutation rate depends on generation time, this effect of population size would be partially cancelled by the generation time effect. Furthermore, the upper limit of heterozygosity is predicted to be due to negative selection.

The slightly deleterious mutation theory is not completely satisfactory, and has been criticized by various authors. The assumption of constant selection coefficients is thought to be unrealistic (*e.g.*, NEI 1987). Thus, we now revise the model by incorporating fluctuation of selection coefficients over space. We

consider this to be realistic for nearly neutral mutations with small effects. In this report, we present results on substitution rates and polymorphisms found from extensive Monte Carlo simulation studies. Some analytical studies will also be included.

NEARLY NEUTRAL MUTATION MODEL

The definition of near neutrality is that the product of effective population size (N) and selection coefficient (s) is not larger than unity. Then the behavior of mutants is influenced by both random drift and selection (KIMURA 1968, 1983). With this kind of small effect, selection coefficients for mutants are assumed to vary among local areas, *i.e.*, spatial fluctuation. This is a new aspect of the nearly neutral model, and comes from the consideration that selective forces are likely to fluctuate depending on both genetic and environmental backgrounds. HARTL and DYKHUIZEN (1985) reported that naturally occurring enzyme variants in *Escherichia coli* usually produced no detectable effects on growth rate under normal conditions, but that many of them produced significant effects when conditions were altered.

The selection coefficient of a new mutant in a local colony is assumed to follow the normal distribution with mean μ_s and variance σ_s^2. Two kinds of population structure are considered; panmictic populations and subdivided populations in a linear habitat with limited migration between adjacent colonies (stepping stone model). In both structures, we assume there are l habitats for the total population, and that each colony occupies a habitat when subdivided. Let $s_{k,i}$ be the selection coefficient for the kth mutant in the ith habitat for a subdivided population. We assume that $s_{k,i}$ are independently chosen from a common normal distribution irrespective of k and i. Thus there is no correlation among $s_{k,i}$'s with respect to k or i. This assumption may not be quite realistic, but for the present purpose of clarifying the effect of fluctuating selection, the model is useful. This kind of fluctuating selection is different from the previous model of varying selection [see HARTL and CLARK (1988), pp. 529–530, for review], in which coefficients fluctuate from one generation to the next. The present model is built to demonstrate long-term effects of various changes of genetic and external backgrounds. Note that there is considerable genetic differentiation among local groups in mammals (NEI 1987) and even in fruit flies (CHOUDHARY and SINGH 1987), and differentiation of the genetic background would influence biochemical processes such as the enzyme kinetics studied in detail by KACSER and BURNS (1981), and hence selective forces.

Our model is also different from the landscape model of GILLESPIE (1983), since we are concerned with the interaction effect of random drift and natural selection, whereas GILLESPIE treated the situation of strong selection and weak mutation. Thus, in GILLESPIE's model, the shift of population occurs by selection and is theoretically tractable. However, in our model, the process of population shift is complicated and the population fitness fluctuates with time because of random drift, as will be shown later. This is the reason we resorted to Monte Carlo simulations.

Let N be the effective population size of a local colony, and $N_T = lN$ be the total size. For panmictic populations, only the total size matters, and the selection coefficient, now denoted by s_k, is assumed to obey the normal distribution with the same mean, μ_s, as before, but with the variance, σ_s^2/l. This is because the mean selection coefficient of l habitats follows this distribution. This model is based on the assumption that the behavior of mutants is determined by the selection coefficients averaged over habitats. The selection coefficient (s_k) of the kth allele in the total population is the average of the selection coefficients ($s_{i,k}$) over habitats and thus the variance of s_k is obtained by dividing σ_s^2 by l. The present study is concerned with the two extreme situations; panmictic populations with average (constant) fitness, and structured populations with the most favorable condition for local differentiation.

The most important parameter is the product of population size and selection coefficient, and $2N\sigma_s$ is assumed to be in the range $0 \sim 1$ (nearly neutral). The value $2N_T\sigma_s$ and $2N_T\sigma_s/\sqrt{l}$ may be considerably larger than unity. The initial population is assumed to be free of mutations. It should be noted that our model differs from the previous models of OHTA (1976), KIMURA (1979) and others, since the distribution of selection coefficients is fixed regardless of the allelic state occupying the population. Let us call it the "fixed model." Note that this is the same as KINGMAN's (1978) house-of-cards model, if random drift is ignored. In contrast, the distribution of selection coefficients is assumed to shift in the previous models, as the population fitness remains unity, even after substitution of mutants. We call this the "shift model." The fixed model corresponds to that adopted by ZENG, TACHIDA and COCKERHAM (1989) and the shift model corresponds to that adopted by HILL (1982) in their studies of the effect of mutation on selection response.

In the shift model, proteins can improve (deteriorate) indefinitely by successive advantageous (deleterious) mutations while their chance of improvement diminishes (increases) as advantageous (deleterious) mutants are fixed in the population in the fixed model. The shift model appears to be unrealistic since there must be a limit for the improvement of proteins. The fixed model also appears unrealistic since any mutant (deleterious or advantageous) can be obtained by a

single step mutation from any allele in this model. Considering the fact that there are many amino acids in a protein, it is not likely that good alleles could be obtained by a single step mutation from any other allele. Thus, both models are unrealistic simplifications and considered to be extreme models between which the real situation lies. Here, we want to incorporate the retardation of substitution as the gene evolves and that is why we adopt the fixed model. In a study of hemoglobin genes, GOODMAN, MOORE and MATSUDA (1975) suggested that evolution had accelerated after gene duplication, but slowed down in later phylogeny, and they argued that this reflects improvement of gene functions. The fixed model would fit this situation.

The average selection coefficient, μ_s, of the present fixed model is the mean selection coefficient of new mutant alleles. Another meaning of μ_s is that it represents the initial condition of the system, since all individuals are free of mutation with fitness 1 at the beginning. The value of μ_s is assumed to be in the range 0 to $-2\sigma_s$. When $\mu_s = 0$, the system starts as "average neutral." For negative μ_s, the system is "slightly deleterious," such that the average selection coefficient of new mutations is negative. Such a situation is expected for proteins or nucleic acids that have existed for a long time and have already attained a desirable function. The average neutral case is for genes newly arisen presumably by duplication, or for old genes but after drastic change of the environment.

SIMULATION EXPERIMENTS

Two sets of Monte Carlo experiments were performed; for subdivided populations and for panmictic populations. A subdivided population consists of l linearly arranged colonies and ten cases of l in the range 1 to 10 were examined. The effective population size of each colony is N, and the total size, $N_T = lN$. The haploid model was used here. Migration was assumed to be between adjacent colonies only with considerably limited rate, m, such that $4Nm = 0.5$. In order to compare the results for subdivided structure with those for random mating populations, ten cases of the population size ($N_T = lN$, $1 \leq l \leq 10$) were examined for panmictic populations.

A model of 100 alleles was adopted, so that every mutation is unique until more than 100 mutations were generated in an experiment for the total population. In practice, mutants are numbered from 1 to 100, and the number is increased by one at each mutation. If the number reaches 100, the mutant is made 1 again. Since highly polymorphic populations are not considered in the present study, new mutants rarely are the same as existing alleles, and hence this model closely approximates the infinite allele model.

The selection coefficient of the kth mutant in the ith colony, $s_{k,i}$, was determined by sampling from a normal distribution with mean μ_s and variance σ_s^2. At the beginning of each experiment, all values of $s_{k,i}$ ($1 \leq k \leq 100$, $1 \leq i \leq l$) were drawn. For panmictic populations, 100 values of s_k were drawn from a normal distribution with mean μ_s and variance σ_s^2/l, when the population size is lN. Sampling and selection were done simultaneously by letting the survival probability of a sampled individual (haploid) equal the fitness, w_i. In practice, the decision on survival is made by using uniform random numbers; if the fitness is less than unity, a single random number determines its survival, and if the fitness is greater than unity, the sampled individual gives at least one progeny and the decision on giving one more progeny is made by a uniform random number.

Because we are concerned with nearly neutral mutations, the value of $2N\sigma_s$ is assumed to be between zero and one. The case of $2N\sigma_s = 1$ was examined in most detail. As for the mutation rate, it was set to treat a segment of DNA that corresponds to an exon or a region smaller than an exon. Let v be the mutation rate for this region of DNA of a gamete per generation. According to KIMURA (1983), the mutation rate per base pair is around 10^{-8} per year. If the long-term effective population size of a typical mammalian species is in the range 10^4 to 10^5, and if an exon contains about 100 base pairs, the product of mutation rate and population size becomes in the range 0.01 to 0.1 for this region, provided that one generation corresponds to 1 year. Thus the value of $Nv = 0.01$ was used in this study. In the simulations, we used $N = 10$ and $v = 0.001$, preserving the value $Nv = 0.01$. Thus N_Tv varies in the range 0.01 to 0.1. Each experiment was continued for 11,000 generations, and the number of mutant substitutions was counted in the whole period. If the mutants are completely neutral, the expected number of substitutions becomes 11 when $v = 0.001$. In practice, when the frequency of a mutant reaches 90% or more, the mutant is counted as it has substituted the previous one.

Heterozygosity and actual number of alleles in a colony and in the total population were recorded in every generation in the range 1,001 to 11,000 generations, and average values were printed. This time interval may not quite represent equilibrium. As will be shown later, the rate of approach to true equilibrium is low, and an enormously long time is required to reach this state, particularly in subdivided populations. Therefore, the results on both the number of substitutions and polymorphism are thought to apply to a transient phase. It should be noted that, if it takes a long time to get to equilibrium, most loci are likely to be in a transient state. Note that the period examined here corresponds to tens of millions of years if

TABLE 1

Number of mutant substitutions in 11,000 generations of the subdivided population

No. of colonies (l)	μ_s:	$\sigma_s = 0.05$ ($2N\sigma_s = 1$)				$\sigma_s = 0.025$ ($2N\sigma_s = 0.5$)		$\sigma_s = 0.01$ ($2N\sigma_s = 0.2$)
		−0.1	−0.05	−0.025	0.0	−0.05	−0.025	−0.01
2		0.98 ± 1.76	3.48 ± 2.15	5.30 ± 2.62	5.53 ± 2.32	5.50 ± 3.42	5.93 ± 2.51	10.63 ± 3.42
3		0.05 ± 0.22	2.55 ± 1.74	3.83 ± 2.00	4.20 ± 2.13	1.53 ± 2.26	5.68 ± 2.82	7.85 ± 2.99
4		0.03 ± 0.16	1.88 ± 1.67	3.40 ± 1.43	3.98 ± 2.08	0.85 ± 1.90	3.53 ± 2.70	6.80 ± 2.75
5		0.00	1.10 ± 1.17	2.43 ± 1.38	2.88 ± 1.54	0.08 ± 0.27	3.08 ± 2.00	6.08 ± 2.47
6		0.00	0.55 ± 0.85	1.98 ± 1.00	2.18 ± 1.13	0.08 ± 0.35	2.55 ± 1.88	4.60 ± 2.16
7		0.00	0.45 ± 0.64	1.25 ± 0.74	1.85 ± 0.92	0.00	1.75 ± 1.46	4.13 ± 1.59
8		0.00	0.30 ± 0.65	1.15 ± 0.83	1.25 ± 0.95	0.00	1.10 ± 1.37	2.95 ± 1.55
9		0.00	0.20 ± 0.46	0.68 ± 0.76	0.85 ± 0.86	0.00	1.10 ± 1.15	2.45 ± 1.26
10		0.00	0.08 ± 0.27	0.60 ± 0.67	0.58 ± 0.55	0.00	0.53 ± 0.68	2.00 ± 1.13

Figures are the mean ± standard deviation of 40 replications. Other parameters: $N = 10$; $v = 10^{-3}$.

time is measured by the unit of $1/v$ generations. Remember that the DNA unit in this study evolves by the rate of 10^{-6} per year, and the period is 11 times the reciprocal of that value. Thus, the equilibrium is unlikely to be attained in the real world. How far the state is from equilibrium depends on the stability of genetic and external environments. Just after gene duplication, or a drastic change of environment, the genetic system is not well adapted, and genes are in a transient state far from equilibrium. However, in a stable gene system, the state may be close to equilibrium. The initial condition, μ_s, represents the situation.

In all experiments, 40 replications were performed for each set of parameters. The number of mutant substitutions and polymorphisms were examined, and will be presented in the next sections.

GENE SUBSTITUTION

Tables 1 and 2 give the results of the number of mutant substitutions in the experiments; Table 1 for the subdivided populations, and Table 2 for the panmictic populations. In Table 2, the substitution numbers under the shift model are also included for comparison. In this model, the distribution of mutant effects is made the same by assigning zero for the selection coefficient of the presently dominant allele. This model is the same as those adopted by OHTA (1976) and KIMURA (1979) except that we employed the normal distribution instead of the gamma distribution (theoretical values for the fixed model are difficult to obtain and cannot be included in the table). When $2N_Tv < 1$, the population is mostly monomorphic experiencing infrequent fixation of mutant alleles. Then the rate of substitution per generation for the shift model, k_s, is the total number of mutants in one generation multiplied by the expected fixation probability (KIMURA 1983) and expressed by,

$$k_s = 2N_Tv \int_{-\infty}^{\infty} f(x) \frac{1 - \exp(-2x)}{1 - \exp(-4N_Tx)}\, dx, \qquad (1)$$

where $f(x)$ is the distribution of the effect of mutation. The values tabulated in Table 2 are computed by numerically integrating this formula and then multiplying this value with the total generation numbers. When the number of substitutions is small, both models behave similarly as can be seen when $\mu_s < -\sigma_s$ in the table. However, if multiple substitutions occur and the population fitness becomes higher than the initial value in absolute sense, a slowdown of the substitution rate ensues in the fixed model whereas no slowdown occurs in the shift model. Thus, the number of substitutions for the shift model is larger than that for the fixed model when $\mu_s > -\sigma_s$, especially when $\mu_s = 0$.

As pointed out before, the expected number of substitutions of completely neutral mutants in the period of our simulations is 11. As can be seen from the tables, the number of substitutions is considerably smaller than the neutral prediction, and it becomes less by increasing the total population size. The result agrees qualitatively with previous models of slightly deleterious mutations (OHTA 1976; KIMURA 1979, 1981; LI 1979, 1987; FOLEY 1987). By examining the results of Tables 1 and 2, it can be said that the effect of population size on mutant substitution is more pronounced for larger value of $2N\sigma_s$. Thus, the negative correlation between substitution rate and population size is expected to be higher for mutations with larger effects under the present model of fluctuating selection.

Another problem is the overly dispersed molecular clock (GILLESPIE 1986, 1987; TAKAHATA 1987), when the variance of the number of amino acid or base substitutions among various species is too large. If the total population size (N_T) as well as environmental changes (μ_s) vary among these species examined, the variance of evolutionary rate is expected to be larger

TABLE 2

Number of mutant substitutions in 11,000 generations of the panmictic population

l ($N_T = lN$) \ μ_s:		$\sigma_s = 0.05/\sqrt{l}$				$\sigma_s = 0.025/\sqrt{l}$		$\sigma_s = 0.01/\sqrt{l}$
		−0.1	−0.05	−0.025	0.0	−0.05	−0.025	−0.01
1	O[a]	3.73 ± 2.82	5.13 ± 2.38	5.70 ± 2.51	6.30 ± 2.68	7.08 ± 3.25	8.23 ± 2.86	10.45 ± 3.46
	E	1.992	5.830	9.166	13.627	4.207	7.289	12.490
2	O	0.60 ± 1.06	3.28 ± 2.16	3.90 ± 2.19	4.95 ± 2.53	3.55 ± 3.80	6.85 ± 2.66	10.00 ± 3.00
	E	0.316	3.200	7.830	16.149	1.372	4.718	10.240
3	O	0.20 ± 0.61	2.25 ± 1.89	3.83 ± 2.49	3.63 ± 2.52	1.65 ± 2.55	5.08 ± 2.82	8.75 ± 3.40
	E	0.048	1.789	6.705	18.327	0.413	3.030	8.334
4	O	0.00	0.93 ± 1.07	3.63 ± 1.88	3.83 ± 1.66	0.25 ± 0.54	3.83 ± 2.77	7.25 ± 2.88
	E	0.007	1.013	5.753	20.276	0.117	1.934	6.735
5	O	0.00	0.60 ± 0.67	2.35 ± 1.39	2.85 ± 1.49	0.00	2.90 ± 2.24	7.63 ± 2.45
	E	0.003	0.579	4.945	22.055	0.043	1.228	5.410
6	O	0.00	0.18 ± 0.45	1.75 ± 0.90	3.35 ± 1.59	0.00	2.10 ± 2.24	7.45 ± 3.55
	E	0.00	0.333	4.257	23.702	0.009	0.777	4.320
7	O	0.00	0.30 ± 0.65	1.70 ± 0.91	2.73 ± 1.80	0.00	0.98 ± 1.48	5.70 ± 2.78
	E	0.00	0.238	3.670	25.244	0.02	0.490	3.435
8	O	0.00	0.28 ± 0.60	1.58 ± 0.96	2.35 ± 1.08	0.00	0.63 ± 1.72	6.10 ± 3.13
	E	0.00	0.129	3.168	26.698	0.00	0.308	2.715
9	O	0.00	0.13 ± 0.33	1.53 ± 0.88	2.23 ± 1.03	0.00	0.10 ± 0.50	5.38 ± 3.18
	E	0.00	0.069	2.737	28.079	0.00	0.193	2.135
10	O	0.00	0.05 ± 0.22	1.48 ± 0.78	2.65 ± 1.33	0.00	0.33 ± 0.76	4.75 ± 3.40
	E	0.00	0.036	2.368	29.395	0.00	0.013	1.675

[a] O, observed (mean ± standard deviation); E, expected, under the shift model. Other parameters: $N = 10$; $v = 10^{-3}$.

than that of the simple Poisson process. Note here that it is quite difficult to distinguish the effect of weak selection from that of the variation of mutation rate itself. It is known that the mutation rate depends much on enzyme system of DNA replication (KUNDEL and ALEXANDER 1986). Since a shift of the enzyme system may occur in evolution, it may have a significant effect on variation of evolutionary rate (BRITTEN 1986). In addition, some genes used by GILLESPIE (1987) are members of multigene families, in which amino acid substitutions are often accelerated, possibly caused by positive selection (OHTA 1988).

POLYMORPHISM

Heterozygosity and actual number of alleles in the total population and in a colony of the subdivided population were studied in the range 1,001 to 11,000 generations of each experiment, and average values over 40 replications were obtained. Tables 3 and 4 present results for subdivided and panmictic populations respectively. Means and standard deviations of heterozygosity and actual numbers of alleles over 40 replications are presented. Note that, under complete neutrality, heterozygosity increases from 0.039 to 0.286, and the actual number changes from 1.14 to 3.14 as $4N_Tv$ moves from 0.04 to 0.4 according to EWENS' (1972) sampling theory by regarding $2N_T$ as the sample size. In the average neutral case, both heterozygosity and the number of alleles in the subdivided, total population are larger than the values expected under strict neutrality, but, heterozygosity and the number of alleles in a local colony do not increase by making larger in the range 4 to 10 colonies. In general, under the slightly deleterious mutation theory, as the population size becomes large, the mutant substitution slows down and heterozygosity approaches the upper limit of mutation-selection balance. This principle holds roughly for the present model.

In each colony, selection works more efficiently in a subdivided structure than in a panmictic population. This can be seen from frequency spectrum of the total population. Figure 1 gives the frequency spectrum of the same simulations as before, that was obtained by averaging 40 replications. Only the case of $l = 2$ is given. The figure shows that one peak exists. When the selection coefficient of a mutant differs between neighboring colonies, the frequency of a mutant stays at intermediate values for a while.

DISCUSSION

The present study is a quantitative evaluation of the previous intuitive argument of the importance of weak selective forces that fluctuate spatially (OHTA 1972b, 1987). The idea is that, in a stable environment, a random mutant needs be beneficial only under re-

TABLE 3

Heterozygosity and actual number of alleles in subdivided populations

l \ μ_s:		$\sigma_s = 0.05$ ($2N\sigma_s = 1$)				$\sigma_s = 0.025$ ($2N\sigma_s = 0.5$)		$\sigma_s = 0.01$ ($2N\sigma_s = 0.2$)
		−0.1	−0.05	−0.025	0.0	−0.05	−0.025	−0.01
Total population								
1	h[a]	0.019 ± 0.006	0.023 ± 0.006	0.023 ± 0.005	0.029 ± 0.005	0.029 ± 0.006	0.030 ± 0.007	0.035 ± 0.008
	n_a	1.078 ± 0.020	1.087 ± 0.019	1.089 ± 0.016	1.090 ± 0.015	1.106 ± 0.018	1.108 ± 0.021	1.122 ± 0.022
2	h	0.039 ± 0.021	0.077 ± 0.031	0.092 ± 0.041	0.082 ± 0.030	0.081 ± 0.025	0.083 ± 0.016	0.103 ± 0.022
	n_a	1.184 ± 0.051	1.285 ± 0.073	1.322 ± 0.097	1.299 ± 0.072	1.307 ± 0.069	1.316 ± 0.044	1.381 ± 0.060
3	h	0.038 ± 0.021	0.139 ± 0.066	0.161 ± 0.067	0.143 ± 0.060	0.081 ± 0.037	0.148 ± 0.041	0.153 ± 0.036
	n_a	1.254 ± 0.060	1.536 ± 0.173	1.599 ± 0.175	1.544 ± 0.158	1.405 ± 0.112	1.594 ± 0.112	1.623 ± 0.103
4	h	0.038 ± 0.015	0.205 ± 0.100	0.275 ± 0.069	0.272 ± 0.089	0.094 ± 0.052	0.190 ± 0.054	0.225 ± 0.038
	n_a	1.326 ± 0.050	1.825 ± 0.259	2.022 ± 0.182	2.004 ± 0.236	1.536 ± 0.159	1.843 ± 0.166	1.969 ± 0.126
5	h	0.038 ± 0.010	0.227 ± 0.102	0.319 ± 0.080	0.325 ± 0.103	0.083 ± 0.022	0.247 ± 0.060	0.299 ± 0.046
	n_a	1.404 ± 0.039	2.026 ± 0.304	2.279 ± 0.236	2.300 ± 0.303	1.614 ± 0.072	2.147 ± 0.207	2.347 ± 0.156
6	h	0.043 ± 0.016	0.255 ± 0.128	0.394 ± 0.098	0.391 ± 0.122	0.094 ± 0.048	0.300 ± 0.086	0.339 ± 0.058
	n_a	1.500 ± 0.070	2.207 ± 0.350	2.674 ± 0.310	2.664 ± 0.360	1.753 ± 0.164	2.463 ± 0.274	2.639 ± 0.218
7	h	0.038 ± 0.011	0.299 ± 0.126	0.419 ± 0.095	0.459 ± 0.097	0.087 ± 0.028	0.327 ± 0.094	0.400 ± 0.061
	n_a	1.550 ± 0.052	2.517 ± 0.405	2.947 ± 0.316	3.086 ± 0.340	1.841 ± 0.111	2.720 ± 0.114	3.034 ± 0.250
8	h	0.039 ± 0.009	0.300 ± 0.122	0.494 ± 0.086	0.535 ± 0.059	0.092 ± 0.023	0.319 ± 0.113	0.434 ± 0.086
	n_a	1.639 ± 0.050	2.684 ± 0.441	3.464 ± 0.362	3.520 ± 0.283	1.968 ± 0.101	2.857 ± 0.406	3.387 ± 0.367
9	h	0.039 ± 0.011	0.298 ± 0.121	0.538 ± 0.081	0.572 ± 0.065	0.088 ± 0.022	0.354 ± 0.126	0.483 ± 0.056
	n_a	1.719 ± 0.065	2.842 ± 0.446	3.790 ± 0.408	3.848 ± 0.349	2.072 ± 0.104	3.176 ± 0.512	3.776 ± 0.304
10	h	0.040 ± 0.010	0.327 ± 0.137	0.566 ± 0.104	0.592 ± 0.066	0.090 ± 0.024	0.421 ± 0.115	0.493 ± 0.059
	n_a	1.801 ± 0.066	3.091 ± 0.527	4.155 ± 0.555	4.193 ± 0.346	2.178 ± 0.122	3.608 ± 0.503	4.028 ± 0.270
Colony								
2	h	0.026 ± 0.010	0.043 ± 0.013	0.050 ± 0.015	0.046 ± 0.014	0.050 ± 0.015	0.052 ± 0.010	0.066 ± 0.014
	n_a	1.103 ± 0.032	1.159 ± 0.044	1.185 ± 0.054	1.168 ± 0.044	1.179 ± 0.046	1.185 ± 0.032	1.228 ± 0.042
3	h	0.024 ± 0.009	0.072 ± 0.026	0.079 ± 0.031	0.077 ± 0.029	0.049 ± 0.018	0.086 ± 0.023	0.094 ± 0.021
	n_a	1.104 ± 0.035	1.265 ± 0.092	1.290 ± 0.105	1.273 ± 0.096	1.185 ± 0.061	1.302 ± 0.074	1.326 ± 0.067
4	h	0.024 ± 0.007	0.085 ± 0.032	0.114 ± 0.034	0.113 ± 0.037	0.051 ± 0.024	0.095 ± 0.027	0.119 ± 0.021
	n_a	1.010 ± 0.024	1.315 ± 0.115	1.406 ± 0.108	1.409 ± 0.130	1.190 ± 0.080	1.337 ± 0.088	1.413 ± 0.068
5	h	0.024 ± 0.005	0.084 ± 0.030	0.110 ± 0.035	0.113 ± 0.037	0.045 ± 0.011	0.102 ± 0.023	0.134 ± 0.021
	n_a	1.100 ± 0.016	1.311 ± 0.109	1.400 ± 0.123	1.413 ± 0.133	1.172 ± 0.036	1.360 ± 0.077	1.466 ± 0.071
6	h	0.024 ± 0.008	0.076 ± 0.029	0.101 ± 0.037	0.117 ± 0.037	0.048 ± 0.014	0.113 ± 0.032	0.131 ± 0.023
	n_a	1.100 ± 0.029	1.282 ± 0.106	1.371 ± 0.126	1.427 ± 0.123	1.178 ± 0.044	1.401 ± 0.107	1.457 ± 0.077
7	h	0.024 ± 0.007	0.092 ± 0.034	0.109 ± 0.034	0.113 ± 0.037	0.047 ± 0.013	0.119 ± 0.031	0.133 ± 0.021
	n_a	1.101 ± 0.022	1.336 ± 0.123	1.403 ± 0.123	1.419 ± 0.133	1.175 ± 0.042	1.418 ± 0.107	1.464 ± 0.071
8	h	0.024 ± 0.007	0.087 ± 0.032	0.128 ± 0.035	0.114 ± 0.029	0.048 ± 0.010	0.101 ± 0.026	0.136 ± 0.023
	n_a	1.102 ± 0.021	1.322 ± 0.109	1.464 ± 0.122	1.418 ± 0.101	1.179 ± 0.033	1.361 ± 0.088	1.474 ± 0.078
9	h	0.023 ± 0.007	0.083 ± 0.025	0.114 ± 0.036	0.108 ± 0.032	0.047 ± 0.012	0.106 ± 0.032	0.145 ± 0.027
	n_a	1.096 ± 0.023	1.305 ± 0.092	1.415 ± 0.126	1.401 ± 0.115	1.175 ± 0.039	1.379 ± 0.110	1.503 ± 0.088
10	h	0.024 ± 0.007	0.082 ± 0.028	0.119 ± 0.041	0.115 ± 0.033	0.047 ± 0.012	0.109 ± 0.029	0.138 ± 0.021
	n_a	1.101 ± 0.022	1.299 ± 0.096	1.434 ± 0.140	1.419 ± 0.112	1.176 ± 0.038	1.385 ± 0.099	1.484 ± 0.072

[a] h, heterozygosity, n_a: actual no. of alleles, for the range 1,001 to 11,000 generations. Figures are the mean ± standard deviation of 40 replications.

stricted circumstances, whereas in a more variable environment, a mutant must be beneficial in many circumstances. Usually, the smaller the population, the more restricted the environmental variability. Based on these considerations, selection coefficients of new mutations are assumed to be normally distributed, *i.e.*, they fluctuate around the mean, μ_s, although for any mutant, the coefficient stays constant over time. The most important parameter, is the variance of the normal distribution. Our model treats nearly neutral mutations, and the product $2N\sigma_s$ is 1 or less. In this regard, it is different from MANI's (1984) model in which selection coefficients fluctuate but the product is assumed to be large.

In several previous attempts to formulate the nearly neutral mutation theory, the selection coefficients of new mutations were assumed to follow a certain distribution such as exponential (OHTA 1976) or gamma (KIMURA 1979). These distributions have no probability density for $s > 0$, *i.e.*, all mutations are slightly

TABLE 4

Heterozygosity and actual number of alleles in panmictic populations

l ($N_T = lN$) \ μ_s:		$\sigma_s = 0.05/\sqrt{l}$				$\sigma_s = 0.025/\sqrt{l}$		$\sigma_s = 0.01/\sqrt{l}$
		−0.1	−0.05	−0.025	0.0	−0.05	−0.025	−0.1
2	h[a]	0.027 ± 0.006	0.037 ± 0.010	0.036 ± 0.012	0.040 ± 0.013	0.042 ± 0.013	0.055 ± 0.011	0.069 ± 0.011
	n_a	1.146 ± 0.020	1.192 ± 0.036	1.188 ± 0.040	1.202 ± 0.040	1.215 ± 0.041	1.250 ± 0.032	1.295 ± 0.035
3	h	0.021 ± 0.005	0.047 ± 0.014	0.054 ± 0.019	0.046 ± 0.015	0.046 ± 0.018	0.074 ± 0.015	0.097 ± 0.020
	n_a	1.202 ± 0.021	1.297 ± 0.047	1.321 ± 0.065	1.296 ± 0.054	1.297 ± 0.060	1.398 ± 0.051	1.468 ± 0.066
4	h	0.019 ± 0.003	0.045 ± 0.011	0.063 ± 0.015	0.061 ± 0.018	0.044 ± 0.012	0.086 ± 0.019	0.119 ± 0.026
	n_a	1.254 ± 0.020	1.376 ± 0.043	1.437 ± 0.061	1.427 ± 0.067	1.378 ± 0.046	1.530 ± 0.068	1.638 ± 0.084
5	h	0.019 ± 0.003	0.046 ± 0.009	0.062 ± 0.019	0.066 ± 0.022	0.040 ± 0.007	0.089 ± 0.024	0.147 ± 0.025
	n_a	1.316 ± 0.026	1.460 ± 0.045	1.525 ± 0.081	1.532 ± 0.083	1.440 ± 0.033	1.636 ± 0.086	1.849 ± 0.088
6	h	0.018 ± 0.002	0.045 ± 0.009	0.067 ± 0.017	0.070 ± 0.026	0.040 ± 0.006	0.095 ± 0.026	0.166 ± 0.032
	n_a	1.368 ± 0.021	1.548 ± 0.048	1.631 ± 0.078	1.638 ± 0.112	1.531 ± 0.041	1.776 ± 0.101	2.022 ± 0.115
7	h	0.018 ± 0.002	0.046 ± 0.009	0.070 ± 0.018	0.079 ± 0.024	0.040 ± 0.006	0.093 ± 0.022	0.177 ± 0.036
	n_a	1.433 ± 0.027	1.624 ± 0.058	1.737 ± 0.083	1.770 ± 0.106	1.618 ± 0.041	1.870 ± 0.088	2.189 ± 0.128
8	h	0.019 ± 0.002	0.051 ± 0.014	0.087 ± 0.024	0.076 ± 0.021	0.040 ± 0.006	0.094 ± 0.029	0.198 ± 0.040
	n_a	1.501 ± 0.027	1.742 ± 0.065	1.897 ± 0.112	1.856 ± 0.089	1.697 ± 0.044	1.978 ± 0.116	2.378 ± 0.153
9	h	0.018 ± 0.002	0.046 ± 0.009	0.094 ± 0.027	0.081 ± 0.026	0.038 ± 0.005	0.083 ± 0.016	0.212 ± 0.048
	n_a	1.550 ± 0.033	1.803 ± 0.045	2.036 ± 0.127	1.982 ± 0.119	1.773 ± 0.044	2.037 ± 0.077	2.566 ± 0.171
10	h	0.019 ± 0.002	0.044 ± 0.007	0.096 ± 0.026	0.104 ± 0.029	0.039 ± 0.005	0.085 ± 0.014	0.212 ± 0.054
	n_a	1.619 ± 0.030	1.881 ± 0.051	2.160 ± 0.136	2.196 ± 0.148	1.862 ± 0.047	2.158 ± 0.070	2.696 ± 0.219

[a] h, heterozygosity; n_a, actual no. of alleles, for the range 1,001 to 11,000 generations. Figures are the mean ± standard deviation of 40 replications.

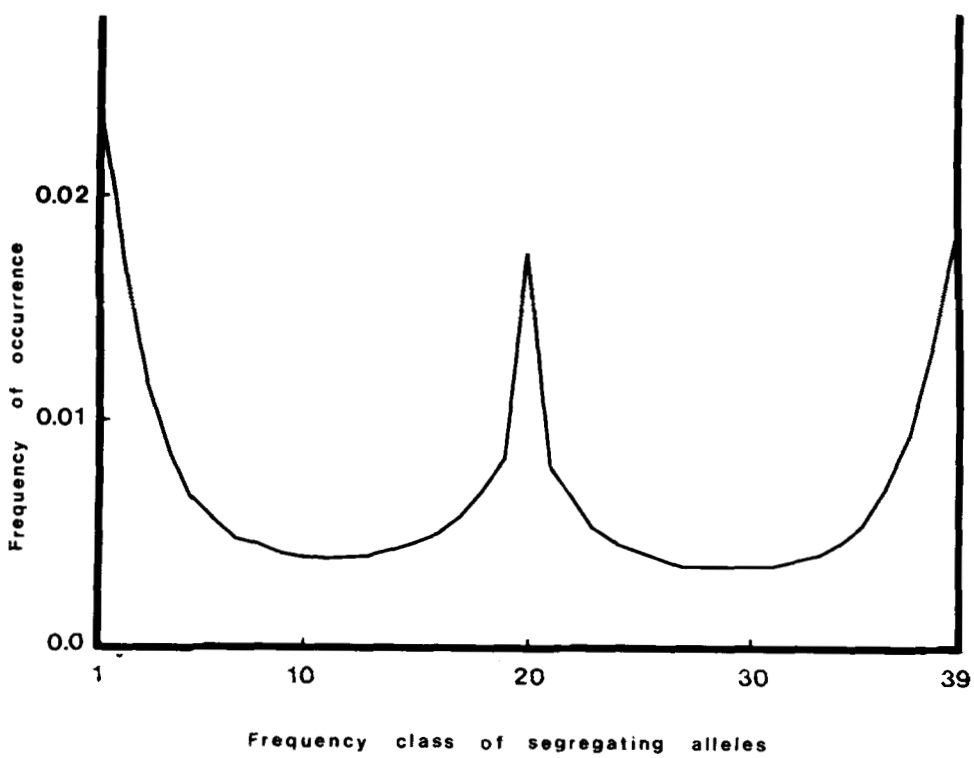

FIGURE 1.—Average frequency spectra of 40 replications for the range 1,001 to 11,000 generations for the case of $2N\sigma_s = 1$, $\mu_s = -\sigma_s$, and $l = 2$.

deleterious. The present model incorporates very slightly advantageous mutations for the region $s > 0$ and is basically different from the previous models. In this regard, our study is an attempt to expand the concept of near neutrality to the realm of the selectionists (*e.g.*, ZUCKERKANDL 1987). Note here that the population moves to a "better" state through selection and drift, and once it has moved, the relative advantage of new mutants becomes negative. However the rate of such "progress" of a population is very low. In Table 5, some typical sample paths for progress of simulated populations are presented in terms of the selection coefficients of successively fixed mutants. For subdivided populations, the values are the averages of $s_{k,i}$ over l colonies ($1 \leq i \leq l$). As can be seen from the table, the selection coefficient varies greatly

TABLE 5

Selection coefficients of successively replaced mutants

Run No.								
Subdivided population								
1	0.0170	−0.0032	0.0452	0.0029	−0.0032	0.0206	0.0390	
2	−0.0210	0.0353	0.0688					
3	0.0368	0.00555	0.0115	−0.0217	0.0383	0.0764	0.0482	0.0764
4	0.0469	0.0071	0.0467	0.0147				
5	0.0064	0.0442	0.0225	0.0415	0.0733			
Panmictic population								
1	0.0241	0.0226	0.0632					
2	0.0671	0.0687	0.0400	0.0600	0.0510			
3	0.0510	0.0668						
4	−0.0166	0.0215	0.0542	0.0553	0.0542			
5	0.0620	0.0660	0.0699	0.0216	0.0393	0.0153	0.0393	0.0699

Five sample paths for each of subdivided population ($2N\sigma_s = 1$, $l = 3$, $\mu_s = 0$) and panmictic population ($\sigma_s = 0.05/\sqrt{3}$, $\mu_s = 0$, $N_T = 30$) are given.

with time, but on the average, it tends to increase.

ZENG *et al.* (1989) have shown that, under the fixed model that is similar to the house of cards model (KINGMAN 1978), selection is very efficient in bringing the equilibrium population fitness to a high value. Using their weak mutation approximation and assuming an infinite number of alleles, the expected selection coefficient in the equilibrium state is shifted upward by $4N_T\sigma_s^2 = \pi\sigma_s$ from the mean (μ_s) of the distribution where $\pi = 4N_T\sigma_s$. Roughly speaking, mutant alleles whose selection coefficients are more than $s_p - 1/(2N_T)$ have significant chances of fixation, where s_p is the selection coefficient of the presently fixed allele. Suppose the expected equilibrium selection coefficient is the selection coefficient of the allele fixed in the equilibrium state. Then in order to have a significant chance of fixation, mutant alleles should have selection coefficients, s_f, with

$$s_f > \left(\pi - \frac{2}{\pi}\right)\sigma_s + \mu_s. \qquad (2)$$

If, for example, $\pi = 4N_T\sigma_s = 4$, they should have selection coefficients greater than $3.5\sigma_s$ plus the mean of the distribution. Such mutants are very rare and evolution is stagnated in this case. However, as noted before, it takes a long time to attain equilibrium and equilibrium may not be realized in nature. That is why we investigated nonequilibrium states in the present paper and why we do not have this stagnation for larger π when μ_s is small in Table 2.

We now examine various observations in the light of our finding. Recent studies on DNA polymorphism in natural populations of Drosophila reveal several facts that are unexpected from knowledge of protein polymorphisms. For example, the level of DNA polymorphism at the *alcohol dehydrogenase* (*Adh*) locus and the region around it seems to be too high compared with that of protein polymorphism (KREITMAN 1987). HUDSON, KREITMAN and Aguade (1987) carried out extensive statistical analyses, and suggested the presence of at least two selected polymorphisms in this region. Such candidates for "balancing" selection could be viewed as an expanded form of the fluctuating selection modeled here, *i.e.*, selective differences between the local environments happens to be large in this particular locus. Even if the migration rate is not as low in *Drosophila melanogaster* as in this study, the frequency spectra in Figure 1 may be attained at *Adh* locus, provided that selection coefficients are larger than the present model.

Another interesting example is the contrasting patterns of DNA polymorphism at the rosy region in *D. melanogaster* and *Drosophila simulans* reported by AQUADRO, LADO and NOON (1988). These authors found that, unlike protein polymorphisms, DNA sequence variation in this region of *D. simulans* is estimated to be several times greater than that of *D. melanogaster*. Note that the average heterozygosity over many protein loci is almost the same in *D. melanogaster* and *D. simulans* but that geographic differentiation is more pronounced in the former than in the latter (CHOUDHARY and SINGH 1987). AQUADRO, LADO and NOON (1988) suggested that differences in species effective population sizes may be responsible for the pattern. AQUADRO (1989) studied more loci, and found the same contrasting pattern in all three loci examined between *D. simulans* and *D. melanogaster*; *per* in the *X* chromosome, *Adh* in the second chromosome in addition to *rosy* in the third chromosome. Although this finding needs to be expanded to more loci in order to be general, it is highly desirable to examine this remarkable pattern from a theoretical standpoint. Our results show that the species effective size as well as population structure influences the pattern of polymorphism when selection intensity fluctuates spatially. Suppose that the important parame-

TABLE 6

Possible sets of parameter values corresponding to the contrasting pattern of protein and DNA polymorphisms

Species	Protein	DNA
D. simulans		
Panmictic, $l = 5$	$\sigma_s = 0.05/\sqrt{5}$	$\sigma_s = 0.01/\sqrt{5}$
	$\mu_s = -0.05$	$\mu_s = -0.01$
	$h = 0.046$	$h = 0.147$
	$n_a = 1.460$	$n_a = 1.849$
D. melanogaster		
Subdivided, $l = 2$	$\sigma_s = 0.05$	$\sigma_s = 0.01$
	$\mu_s = -0.05$	$\mu_s = -0.01$
	Total population	
	$h = 0.077$	$h = 0.103$
	$n_a = 1.285$	$n_a = 1.381$
	Colony	
	$h = 0.043$	$h = 0.066$
	$n_a = 1.159$	$n_a = 1.228$

ter, $2N\sigma_s$, is larger in protein loci than in loci detectable by DNA analysis. The extent of polymorphism may be much different, if the population of a species is large and panmictic, and that of another species is not so large and subdivided (CHOUDHARY and SINGH 1987). For example, suppose that the following sets of parameter values correspond roughly to the values of the two species. Then the extent of polymorphisms would be as in Table 6, by using data of Tables 3 and 4. Then the contrasting pattern may be understood. Of course, the comparison of the theoretical prediction and actual data merely indicates the direction of the difference of the parameter values between the two *Drosophila* species, and more detailed examination is needed for exact understanding.

LANGLEY and his associates extensively studied DNA polymorphisms at the white-locus region of *D. melanogaster* (LANGLEY and AQUADRO 1987; MIYASHITA and LANGLEY 1988). They found that there is no difference in the level of DNA polymorphisms between loci on autosomes and on the *X* chromosome. From this, they argue that negative selection would be uncommon for molecular variation. However, if the selection is so mild that $2N\sigma_s$ is unity or less as studied here, it would not matter whether the gene is single-dose (*X* chromosome) or double-dose (autosomal) since there is almost no dominance for mutations with small effects (MUKAI and YAMAGUCHI 1974; KASCER and BURNS 1981).

Another important subject related to our study is that of the rate of molecular evolution. The molecular clock has been controversial for more than 20 years involving both population geneticists and molecular evolutionists (see a collection of papers in the special issue of *Journal of Molecular Evolution*; JUKES 1987). One problem is the dependency on generation time (LAIRD, MCCONAUGHY and MCCARTHY 1969; KOHNE 1970; WU and LI 1985; BRITTEN 1986; MORIYAMA 1987; SIBLEY and AHLQUIST 1987). In our nearly neutral mutation model, as well as in the strictly neutral model, the evolutionary rate is highly dependent on mutation rate (usually measured per cell division) [see ALBERTS *et al.* (1989), p. 96]. A recent report on male-driven evolutionary rate (MIYATA *et al.* 1987) also suggests that evolutionary rate is determined by the number of germ cell cycles. In mammals, both male and female primordial germ cells divide early in embryogenesis, and are stored until sexual maturation. At maturation, only in males, the cells begin to proliferate, attaining an ability to divide indefinitely [ALBERTS *et al.* (1989), pp. 862–865]. Thus, if the cell cycle of the male germ cell line is important, the evolutionary rate would be *mildly* dependent on maturation time and hence on individual generation time. Data on DNA hybridization and synonymous changes strongly indicate that animals with short generation time evolve rapidly (for reviews, see LI, LUO and WU 1985; SHARP 1989). However, the rate of amino acid substitution seems to be less dependent on generation time (LI, TANIMURA and SHARP 1987; GILLESPIE 1989).

Our results show that the effect of total population size on the substitution rate is larger when $2N\sigma_s = 0.5$ and 1 than when $2N\sigma_s = 0.2$, indicating that the cancellation between generation time effect and negative selection would be more pronounced for amino acid substitutions than synonymous changes. Data of LI, TANIMURA and SHARP (1987) suggest that the generation time effect is stronger on synonymous substitutions than on amino acid replacement substitutions, agreeing with our results. By using data of LI, TANIMURA and SHARP (1987), GILLESPIE (1989) obtained weighting factors, which are the ratios of the numbers of substitutions of the three branches for each mammalian orders, artiodactyls, rodents and primates. Let w_a, w_r and w_p be the weights for the above three orders, respectively. For replacement substitutions, they turned out to be

$$w_a = 0.885,\ w_r = 1.279,\ w_p = 0.836.$$

For the synonymous substitutions,

$$w_a = 0.762,\ w_r = 1.611,\ w_p = 0.627.$$

Thus, amino acid replacement substitutions are less dependent on generation time.

As repeatedly emphasized, the generation time effect is coupled with the population size effect. An interesting example that shows the population size effect is the differentiation of mitochondrial genome among Hawaiian Drosophila reported by DESALLE and TEMPLETON (1988). These authors examined the rate of mitochondrial evolution in two closely related lineages of Hawaiian Drosophila that have different

histories, and found that the rate is three times higher in lineages with repeated founder events than in lineages without bottle necks. Thus mitochondrial evolution fits the prediction of our model.

The final problem is concerned with the assumption of fluctuating selection coefficients. Results of DYKHUIZEN and HARTL (1980) suggest that many naturally occurring enzyme polymorphisms in *E. coli.* are neutral or nearly neutral, but a latent potential for selection can be observed in the polymorphism at the 6-phosphogluconate dehydrogenase locus of *E. coli*, SAWYER, DYKHUIZEN and HARTL (1987) estimated that the average selection coefficient for that locus is approximately -1.6×10^{-7}. This estimate corresponds to our μ_s, and local values of selection coefficient may be larger. DEAN, DYKHUIZEN and HARTL (1988) examined the fitness effect of newly obtained amino acid substitutions at the β-galactosidase locus of *E. coli*, and again found that the majority of amino acid changes have minor effects on fitness. This would imply that the "Dykhuizen-Hartl" effect applies not only to naturally occurring polymorphic alleles but also to many newly arisen mutant alleles. DEAN, DYKHUIZEN and HARTL (1988) also argue that the amino acid replacements have negligible effects on fitness in their lactose-limited chemostats, but that they may well have slightly deleterious effects in natural populations.

We thank B. S. WEIR and two anonymous referees for their very useful comments on the manuscript. This work is supported by a grant-in-aid from the Ministry of Education, Science and Culture of Japan. Contribution no. 1824 from the National Institute of Genetics, Mishima, 411 Japan.

LITERATURE CITED

ALBERTS, B., D. BRAY, J. LEWIS, M. RAFF, K. ROBERTS and J. D. WATSON, 1989 *Molecular Biology of the Cell*, Ed. 2. Garland Publishing, New York.

AQUADRO, C. F., 1989 Contrasting levels of DNA sequence variation in *Drosophila* species revealed by "six-cutter" restriction map surveys. UCLA Symp. Mol. Cell. Biol. **122:** (in press).

AQUADRO C. F., K. M. LADO and W. A. NOON, 1988 The rosy region of *Drosophila melanogaster* and *Drosophila simulans*. I. Contrasting levels of naturally occurring DNA restriction map variation and divergence. Genetics **119:** 875–888.

BRITTEN, R. J., 1986 Rates of DNA sequence evolution differ between taxonomic groups. Science **231:** 1393–1398.

CHOUDHARY, M., and R. S. SINGH, 1987 A comprehensive study of genetic variation in natural populations of *Drosophila melanogaster*. III. Variations in genetic structure and their causes between *Drosophila melanogaster* and its sibling species *Drosophila simulans*. Genetics **117:** 697–710.

DEAN, A. M., D. E. DYKHUIZEN and D. L. HARTL, 1988 Fitness effects of amino acid replacements in the β-galactosidase of *Escherichia coli*. Mol. Biol. Evol. **5:** 469–485.

DESALLE, R., and A. R. TEMPLETON, 1988 Founder effects and the rate of mitochondrial DNA evolution in Hawaiian *Drosophila*. Evolution **42:** 1076–1084.

DYKHUIZEN, D. E., and D. L. HARTL, 1980 Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. Genetics **96:** 801–817.

EWENS, W. J., 1972 The sampling theory of selectively neutral alleles. Theor. Popul. Biol. **3:** 87–112.

FOLEY, P., 1987 Molecular clock at loci under stabilizing selection. Proc. Natl. Acad. Sci. USA **84:** 7996–8000.

GILLESPIE, J. H., 1983 A simple stochastic gene substitution model. Theor. Popul. Biol. **23:** 202–215.

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. Oxf. Surv. Evol. Biol. **4:** 10–37.

GILLESPIE, J. H., 1989 Lineage effects and the index of dispersion of molecular evolution. Mol. Biol. Evol. **6:** 636–647.

GOODMAN, M., G. W. MOORE and G. MATSUDA, 1975 Darwinian evolution in the genealogy of hemoglobin. Nature **253:** 603–608.

HARTL, D. L., and A. G. CLARK, 1988 *Principles of Population Genetics*, Ed. 2. Sinauer Associates, Sunderland, Mass.

HARTL, D. L., and D. E. DYKHUIZEN, 1985 The neutral theory and the molecular basis of preadaptation, pp. 107–124 in *Population Genetics and Molecular Evolution*, edited by T. OHTA and K. AOKI. Japan Scientific Societies Press, Tokyo.

HILL, W. G., 1982 Rates of change in quantitative traits from fixation of new mutations. Proc. Natl. Acad. Sci. USA **79:** 142–145.

HUDSON, R. R., M. KREITMAN and M. AGUADE, 1987 A test of neutral molecular evolution based on nucleotide data. Genetics **116:** 153–159.

JUKES, T. H. (Editor), 1987 Special issue–molecular evolutionary clock. J. Mol. Evol. **26:** nos. 1 and 2.

KACSER, H., and J. A. BURNS, 1981 The molecular basis of dominance. Genetics **97:** 639–666.

KIMURA, M., 1968 Evolutionary rate at the molecular level. Nature **217:** 624–626.

KIMURA, M., 1979 A model of effectively neutral mutations in which selective constraint is incorporated. Proc. Natl. Acad. Sci. USA **76:** 3440–3444.

KIMURA, M., 1981 Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. Proc. Natl. Acad. Sci. USA **78:** 5773–5777.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, New York.

KINGMAN, J. F. C., 1978 A simple model for the balance between selection and mutation. J. Appl. Probab. **15:** 1–12.

KOHNE, D. E., 1970 Evolution of higher-organism DNA. Q. Rev. Biophys. **3:** 327–375.

KREITMAN, M., 1987 Molecular population genetics. Oxf. Surv. Evol. Biol. **4:** 38–60.

KUNKEL, T. A., and P. S. ALEXANDER, 1986 The base substitution fidelity of eukaryotic DNA polymerases. Mispairing frequencies site preferences, insertion preferences and base substitution by dislocation. J. Biol. Chem. **261:** 160–166.

LAIRD, C. D., B. L. MCCONAUGHY and B. J. MCCARTHY, 1969 Rate of fixation of nucleotide substitutions in evolution. Nature **224:** 149–154.

LANGLEY, C. H., and C. F. AQUADRO, 1987 Restriction map variation in natural populations of *Drosophila melanogaster*: *white* locus region. Mol. Biol. Evol. **4:** 651–663.

LI, W.-H., 1979 Maintenance of genetic variability under the pressure of neutral and deleterious mutations in a finite population. Genetics **92:** 647–667.

LI, W.-H., 1987 Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. J. Mol. Evol. **24:** 337–345.

LI, W.-H., C. C. LUO and C.-I. WU, 1985 Evolution of DNA sequences, pp. 1–94 in *Molecular Evolutionary Genetics*, edited by R. J. MACINTYRE. Plenum Press, New York.

LI, W.-H., M. TANIMURA and P. M. SHARP, 1987 An evaluation

of the molecular clock hypothesis using mammalian DNA sequences. J. Mol. Evol. **25:** 330–342.

MANI, G. S., 1984 A Darwinian theory of enzyme polymorphism, pp. 242–298 in *Evolutionary Dynamics of Genetic Diversity* (Lecture Notes in Biomathematics, Vol. 53), edited by G. S. MANI. Springer, Berlin.

MIYASHITA, N., and C. H. LANGLEY, 1988 Molecular and phenotypic variation of the *white* locus region in *Drosophila melanogaster*. Genetics **120:** 199–212.

MIYATA, T., H. HAYASHIDA, K. KUMA, K. MITSUYA and T. YASUNAGA, 1987 Male-driven molecular evolution: a model and nucleotide sequence analysis. Cold Spring Harbor Symp. Quant. Biol. **52:** 863–867.

MORIYAMA, E. N., 1987 Higher rates of nucleotide substitution in *Drosophila* than in mammals. Jpn. J. Genet. **62:** 139–147.

MUKAI, T., and O. YAMAGUCHI, 1974 The genetic structure of natural populations of *Drosophila melanogaster*. XI. Genetic variability of local populations. Genetics **76:** 339–366.

NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

OHTA, T., 1972a Evolutionary rate of cistrons and DNA divergence. J. Mol. Evol. **1:** 150–157.

OHTA, T., 1972b Population size and rate of evolution. J. Mol. Evol. **1:** 305–314.

OHTA, T., 1973 Slightly deleterious mutant substitutions in evolution. Nature **246:** 96–98.

OHTA, T., 1974 Mutational pressure as the main cause of molecular evolution and polymorphisms. Nature **252:** 351–354.

OHTA, T., 1976 Role of very slightly deleterious mutations in molecular evolution and polymorphism. Theor. Popul. Biol. **10:** 254–275.

OHTA, T., 1977 Extension to the neutral mutation random drift hypothesis, pp. 148–167 in *Molecular Evolution and Polymorphism*, edited by M. KIMURA. National Institute of Genetics, Mishima.

OHTA, T., 1987 Very slightly deleterious mutations and the molecular clock. J. Mol. Evol. **26:** 1–6.

OHTA, T., 1988 Multigene and supergene families. Oxf. Surv. Evol. Biol. **5:** 41–65.

SAWYER, S. A., D. E. DYKHUIZEN and D. L. HARTL, 1987 Confidence interval for the number of selectively neutral amino acid polymorphisms. Proc. Natl. Acad. Sci. USA **84:** 6225–6228.

SHARP, P. M., 1989 Evolution at 'silent' sites in DNA, pp. 23–32 in *Evolution and Animal Breeding*, edited by W. G. HILL and T. F. C. MACKEY. C. A. B. INTERNATIONAL, WALLINGTON, U.K.

SIBLEY, C. G., and J. E. AHLQUIST, 1987 DNA hybridization evidence of hominoid phylogeny: results from an expanded data set. J. Mol. Evol. **26:** 99–121.

TAKAHATA, N., 1987 On the overdispersed molecular clock. Genetics **116:** 169–179.

WU, C.-I., and W.-H. LI, 1985 Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA **82:** 1741–1745.

ZENG, Z.-B., H. TACHIDA and C. C. COCKERHAM, 1989 Effects of mutation on selection limits in finite populations with multiple alleles. Genetics **122:** 977–984.

ZUCKERKANDL, E., 1987 On the molecular evolutionary clock. J. Mol. Evol. **26:** 34–46.

Communicating editor: B. S. WEIR