

Evolutionarily Stable Mutation Rate in a Periodically Changing Environment

Kazushige Ishii,* Hirotugu Matsuda,[†] Yoh Iwasa[†] and Akira Sasaki[†]

*College of General Education, Nagoya University, Nagoya 464, Japan and [†]Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812, Japan

Manuscript received December 9, 1987

Accepted for publication September 23, 1988

ABSTRACT

Evolution of mutation rate controlled by a neutral modifier is studied for a locus with two alleles under temporally fluctuating selection pressure. A general formula is derived to calculate the evolutionarily stable mutation rate μ_{ess} in an infinitely large haploid population, and following results are obtained. (I) For any fluctuation, periodic or random: (1) if the recombination rate r per generation between the modifier and the main locus is 0, μ_{ess} is the same as the optimal mutation rate μ_{op} which maximizes the long-term geometric average of population fitness; and (2) for any r , if the strength s of selection per generation is very large, μ_{ess} is equal to the reciprocal of the average number τ of generations (duration time) during which one allele is persistently favored than the other. (II) For a periodic fluctuation in the limit of small s and r , $\mu_{\text{ess}}\tau$ is a function of $s\tau$ and $r\tau$ with properties: (1) for a given $s\tau$, $\mu_{\text{ess}}\tau$ decreases with increasing $r\tau$; (2) for $s\tau \leq 1$, $\mu_{\text{ess}}\tau$ is almost independent of $s\tau$, and depends on $r\tau$ as $\mu_{\text{ess}}\tau \approx 1.6$ for $r\tau \ll 1$ and $\mu_{\text{ess}}\tau \approx 6/r\tau$ for $r\tau \gg 1$; and (3) for $s\tau \geq 1$, and for a given $r\tau$, $\mu_{\text{ess}}\tau$ decreases with increasing $s\tau$ to a certain minimum less than 1, and then increases to 1 asymptotically in the limit of large $s\tau$. (III) For a fluctuation consisting of multiple Fourier components (*i.e.*, sine wave components), the component with the longest period is the most effective in determining μ_{ess} (*low pass filter effect*). (IV) When the cost c of preventing mutation is positive, the modifier is non-neutral, and μ_{ess} becomes larger than in the case of neutral modifier under the same selection pressure acting at the main locus. The value of c which makes μ_{ess} equal to μ_{op} of the neutral modifier case is calculated. It is argued that this value gives a critical cost such that, so long as the actual cost exceeds this value, the evolution rate at the main locus must be smaller than its mutation rate μ_{ess} .

IN a constant environment, mutation is deleterious since it brings about a mutational load to the population by producing unfit alleles from a common one which is often best fit to the environment. Theoretical studies of mutation rate modifier dynamics conclude that the rate should evolve toward zero (for example, LIEBERMAN and FELDMAN 1986). In a fluctuating environment, however, the rate can evolve toward a nonzero level since mutation is advantageous as long as it provides genetic variation necessary for a population to adapt to the changing environment.

This idea was put forward by STURTEVANT (1937), and was quantitatively studied by KIMURA (1960, 1967). KIMURA proposed that mutation rate would evolve toward a rate which minimizes the sum of the mutation load L_m and the substitution load L_e . LEVINS (1967) further developed this idea by calculating the optimal mutation rate μ_{op} which maximizes the long-term geometric average of population fitness for a model that explicitly incorporated a fluctuating environment. Recently, ISHII and MATSUDA (1985) proved that the optimal mutation rate μ_{op} is equal to the evolution rate v which is defined as the rate of mutant substitutions that have occurred along the phylogenetic line leading to the present organisms.

However, all of these papers implicitly assumed that the mutation rate is adjusted through group selection. So, their conclusions need to be compared with those of modifier theories in which the mutation rate evolves through individual selection. This is because a population which is best fit with respect to group selection can sometimes be unstable against the invasion of a mutant modifier through individual selection.

LEIGH (1970, 1973) studied for the first time a mutation modifier model in fluctuating environment. For a special case of very strong selection, he showed that a nonzero optimal level is realized by neutral modifiers in an asexual population. He further argued that in a sexual population selection would adjust the mutation rate toward zero, but did not explicitly calculate the level of mutation rate attained by modifiers. GILLESPIE (1981) studied a modifier model similar to LEIGH's but in a rapidly fluctuating environment which can be analyzed by a diffusion model. In a case of very loosely linked modifiers, he found that, depending on the model parameters, selection on mutation rates will operate toward three goals: the highest possible rate, the lowest possible rate, or an inter-

mediate rate. This result is qualitatively different from LEIGH's.

In this paper we study a population genetic model of mutation rate modifiers in an infinitely large haploid population. We derive a formula by which we can calculate the evolutionarily stable mutation rate μ_{ess} which is to be realized as an evolutionary consequence. We examine how μ_{ess} depends on the strength and duration of periodically fluctuating selection and the recombination between the modifier and the main locus. We find that selection would generally adjust μ_{ess} at a nonzero level even in the weak selection limit or by unlinked modifiers. We discuss possible differences of μ_{ess} between a periodic environment and a random one. Finally, we calculate the effect of non-neutral modifiers for a case of positive cost of preventing mutation, and discuss its implication on the molecular evolution rate at the main locus.

MODIFIER MODEL

We consider a two-locus model of mutation rate modifiers in a haploid population of effectively infinitely large size. The main locus with two alleles A and a is under a fluctuating selection such that the relative fitnesses of A and a in the t th generation are $1 + s(t)$ and $1 - s(t)$, respectively. The selection coefficient $s(t)$ fluctuates through time with the average 0. The modifier locus with two alleles B and b controls the mutation rate between A and a at the main locus as μ and μ' for B and b , respectively. The recombination rate between the modifier and the main locus is r per generation. The modifier alleles are selectively neutral so that the fitness of a genome does not depend on them.

Consider a population which is made up of only B -carrying genomes. We introduce a small fraction of mutant modifier b carrying genomes with a different mutation rate μ' into it, and ask if the mutant modifier increases in the population or not. If no mutant modifier b in a given set of modifiers can increase in the population, the resident modifier B is said to be evolutionarily stable (MAYNARD SMITH and PRICE 1973), and we call the mutation rate μ caused by B an evolutionarily stable mutation rate μ_{ess} . As a consequence of repeated introduction of new modifiers, we expect that the mutation rate would evolve toward μ_{ess} .

We can analyze the elimination of an introduced b modifier by a following linear dynamical system model since b carrying genomes can be assumed to remain rare in the population throughout the course. Let $N_1(t)$ and $N_2(t)$ be, respectively, the numbers of Ab carrying genomes and ab carrying ones in the population at time t . Then, the corresponding numbers at the next

time $t + 1$ are obtained in three steps as follows. First, genic selection amplifies them as

$$N_1' = (1 + s(t))N_1(t), \quad N_2' = (1 - s(t))N_2(t).$$

Next, mutation with rate μ' modifies them as

$$\begin{aligned} N_1'' &= (1 - \mu')N_1' + \mu'N_2', \\ N_2'' &= \mu'N_1' + (1 - \mu')N_2'. \end{aligned}$$

Finally, recombination completes the change in one generation as

$$\begin{aligned} N_1(t + 1) &= \{(1 - r) + rx_1(t)\}N_1'' + rx_1(t)N_2'', \\ N_2(t + 1) &= rx_2(t)N_1'' + \{(1 - r) + rx_2(t)\}N_2''. \end{aligned}$$

Here, $x_1(t)$ and $x_2(t)$ are, respectively, the frequencies of A and a among B carrying genomes just before recombination takes place. Recombination with b carrying genomes is neglected since they are rare in the population.

Combining above three steps, we obtain the numbers of b carrying genomes at time $t + 1$ from those at time t as

$$\begin{bmatrix} N_1(t + 1) \\ N_2(t + 1) \end{bmatrix} = M(t) \begin{bmatrix} N_1(t) \\ N_2(t) \end{bmatrix}, \quad (1)$$

where 2×2 matrix $M(t)$ is given by

$$\begin{aligned} M(t) = & \begin{bmatrix} \{1 - rx_2(t) - (1 - r)\mu'\}\{1 + s(t)\} \\ \{rx_2(t) + (1 - r)\mu'\}\{1 + s(t)\} \\ \{rx_1(t) + (1 - r)\mu'\}\{1 - s(t)\} \\ \{1 - rx_1(t) - (1 - r)\mu'\}\{1 - s(t)\} \end{bmatrix}. \end{aligned} \quad (2)$$

In order to apply (1) and (2) to study the ultimate fate of b carrying genomes, we must first know values of frequencies $x_1(t)$ and $x_2(t) = 1 - x_1(t)$ of A and a among B carrying genomes at each time t . When the population consists of only B genomes, their time change is due to fluctuating selection and mutation with rate μ , and is given by

$$x_1(t + 1) = \frac{(1 - \mu)\{1 + s(t + 1)\}x_1(t) + \mu\{1 - s(t + 1)\}\{1 - x_1(t)\}}{1 + s(t + 1)\{2x_1(t) - 1\}}. \quad (3)$$

This applies also after the introduction of b carrying genomes as long as they remain rare in the population, because recombination with them can be neglected.

According to (1), the long-term increase rate

$$\lambda = \lim_{i \rightarrow \infty} \{N(t)/N(0)\}^{1/i}$$

of the total number $N(t) = N_1(t) + N_2(t)$ of b carrying genomes is determined by the multiplication of matrices $\{M(t); t = 0, 1, 2, \dots\}$. Since the matrix $M(t)$ in (2) depends not only on μ' but also on μ through $x_i(t)$ determined by (3), the rate λ is a function of both μ' and μ , say $\lambda(\mu', \mu)$. Then, in order that μ be stable

against the introduction of modifier alleles with a slightly different μ' , $\lambda(\mu', \mu)$ must be smaller than $\lambda(\mu, \mu)$. This gives

$$\Lambda(\mu) = [\partial\lambda(\mu', \mu)/\partial\mu']_{\mu'=\mu} = 0 \quad (4a)$$

together with

$$\Lambda(\mu') \geq 0 \quad \text{for} \quad \mu \geq \mu' \quad (4b)$$

as the condition for μ to be an evolutionarily stable mutation rate μ_{ess} .

The long-term increase rate λ is equal to the long-term geometric average $\lim_{t \rightarrow \infty} \{\bar{w}(t-1)\bar{w}(t-2) \dots \bar{w}(0)\}^{1/t}$ of

$$\bar{w}(t) = [\{1 + s(t)\}N_1(t) + \{1 - s(t)\}N_2(t)]/N(t)$$

since $N(t) = \bar{w}(t-1)\bar{w}(t-2) \dots \bar{w}(0)N(0)$ according to (1)–(2). Here, $\bar{w}(t)$ is the average fitness of a subpopulation of b carrying genomes. It is a function of not only μ' but also μ as long as $M(t)$ in (2) depends on μ . In a particular case where the modifier locus is completely linked with the main locus ($r = 0$), $M(t)$ does not depend on μ , hence λ is independent of μ and is equal to the long-term geometric average of population fitness for the case of a homogeneous modifier locus with mutation rate μ' . Therefore we can conclude by (4) that *for completely linked modifiers, μ_{ess} is the same as the optimal mutation rate μ_{op} which maximizes the long-term geometric average of population fitness.*

However, in general cases, some recombination may occur each generation between the modifier and the main locus ($r > 0$). Then the optimal mutation rate μ_{op} may not be attained as a consequence of natural selection. In order to study more clearly what happens in this case, we concentrate in the following on the periodically fluctuating environment with a finite period T . Possible differences in results between a periodic environment and a random one will be discussed later.

In a periodic environment with period T , the selection coefficient $s(t)$ is a periodic sequence with period T . Then, since the periodic transformation (3) with period T has been exerted for a long time, time sequences of $x_1(t)$ and $x_2(t)$ must have converged to periodic ones with the same period T by the time of introduction of b carrying genomes,¹ and remain so as long as b carrying genomes are rare in the population. The periodic sequence of $x_1(t)$ is obtained by solving simultaneous equations (3) for $t = 0, 1, 2, \dots, T-1$ under a boundary condition $x_1(T) = x_1(0)$.

Since we now know that time sequences of $x_1(t)$ and

$x_2(t)$ are periodic with period T , we see that the time sequence of matrix $M(t)$ in (2) is also periodic with the same period T . Thus, the iteration of (1) is much simplified as

$$\begin{bmatrix} N_1(nT) \\ N_2(nT) \end{bmatrix} = M_T^n \begin{bmatrix} N_1(0) \\ N_2(0) \end{bmatrix} \quad (n = 0, 1, 2, \dots), \quad (5)$$

where M_T is a constant matrix given by

$$M_T = M(T-1)M(T-2) \dots M(2)M(1)M(0). \quad (6)$$

According to (5), the long-term increase rate λ of numbers of b carrying genomes is given by the T th root of the greatest eigenvalue of matrix M_T . Therefore, μ_{ess} is determined by (4) if we use this eigenvalue as $\lambda(\mu', \mu)$. Although analytical solution of (4) is limited to specially simple cases, its numerical solution is easy for any periodic selection and recombination rate r as is explained in APPENDIX A.

Note that to obtain μ_{ess} by numerically solving (4) is quite different from finding μ_{ess} based on a computer simulation of modifier competitions. We numerically solve (4) by a routine of bisection using $\Lambda(\mu)$ values calculated by multiplying 2×2 matrices T times and by solving quadratic equations. By this method we can obtain the precise value of μ_{ess} in a short computation time for any model parameters. Arbitrarily weak selection and long period T as large as 10^7 , which are practically intractable by computer simulations, are no problem to our method.

RESULTS

Let us consider a periodic selection with period $T = 2\tau$ where $s(t)$ is given as

$$s(t) = \begin{cases} +s & \text{for } t = 0, 1, 2, \dots, \tau - 1 \\ -s & \text{for } t = \tau, \tau + 1, \dots, T - 1. \end{cases} \quad (7)$$

Here, s is a positive parameter to denote the strength of selection and τ is the number of generations during which the same selection pressure continues to work. We are interested in how μ_{ess} depends on the strength s and the duration τ of selection, and the recombination rate r between the modifier and the main locus.

Figure 1 shows μ_{ess} vs. r for different $\tau = 1, 2, 3, 4, 5, 10, 25$ and $s = 0.05, 0.5$. For a given pair of s and τ , μ_{ess} generally decreases monotonically as r increases from 0 (complete linkage) to 0.5 (free recombination). However, as long as $r\tau \leq 1$, the decrease with r is not very large: even for $r = 0.5$ and $\tau = 5$ ($r\tau = 2.5$), μ_{ess} is more than 75% of μ_{op} which equals μ_{ess} for $r = 0$. The decrease becomes appreciable only when $r\tau \gg 1$.

Several exceptions from the monotone decrease of μ_{ess} with r are seen in Figure 1. First, for $\tau = 1$ and 2, we see that μ_{ess} does not depend on r . For any value of s and r , μ_{ess} is always found at 1 for $\tau = 1$ and at 0.5 for $\tau = 2$. This result is confirmed analytically in

¹ The convergence to periodic sequences with period T can be proved by noting that apparently nonlinear transformation (3) of the frequency $x_1(t)$ is related to a periodic linear transformation of genome numbers like (1)–(2) but with $r = 0$ and μ' replaced by μ .

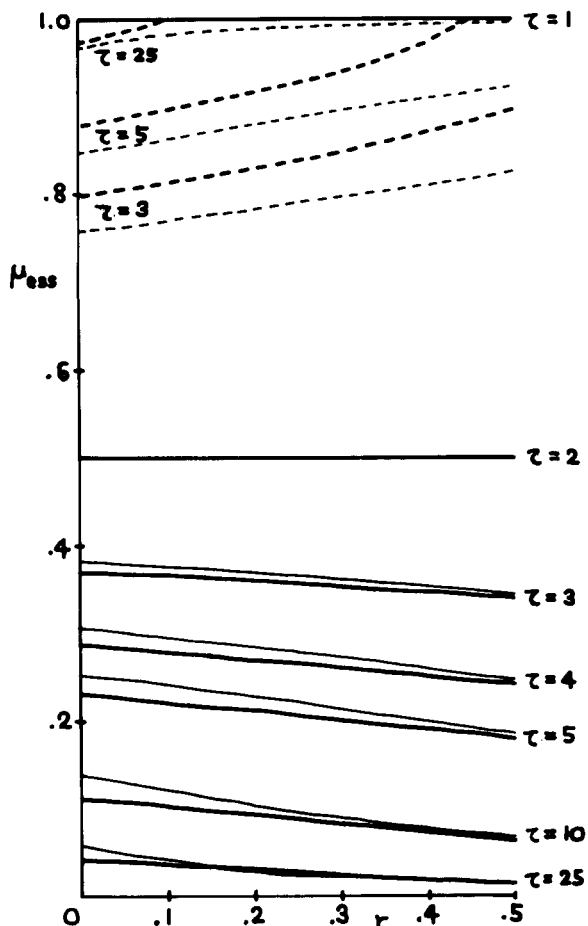


FIGURE 1.—The ESS mutation rate μ_{ess} vs. the recombination rate r between the modifier and the main loci for a periodic selection with a single strong Fourier component (7). Curve is drawn for a given pair of strength s and duration τ of selection. Thick lines are for $s = 0.5$ and thin ones for $s = 0.05$. τ is chosen from 1, 2, 3, 4, 5, 10 and 25. For $\tau = 3, 5$ and 25, broken lines give the demarcation mutation rate which bounds the attractor of the second ESS mutation rate $\mu_{\text{ess}} = 1$ not shown in the figure.

the weak selection limit $s\tau \ll 1$ (see APPENDIX B) and if the strength of selection is very large $s \approx 1$ (see APPENDIX C).

Next, for odd numbers of τ , we see that there are two μ_{ess} . The larger one is always at 1 for any s and r , while the smaller one changes with s and r . So, the range $[0, 1]$ of μ is divided into two attractors of each μ_{ess} , and the demarcation mutation rate which separates them is also shown in Figure 1 by broken lines.

We may not need worry about $\mu_{\text{ess}} = 1$ for odd τ case with $\tau \geq 3$, because its attractor extends only in a too high region of mutation rate ($\mu > 0.7$) to allow a biologically meaningful interpretation. However, $\mu_{\text{ess}} = 1$ and 0.5, respectively, for $\tau = 1$ and 2 may be interpreted as to suggest that the mutation rate would evolve toward the highest possible level in such a rapidly oscillating environment.

When s , μ and r are small, the behavior of our discrete time model is expected to be approximated reasonably well by a continuous time model. Then,

μ_{ess} will depend on parameters s , τ and r in such a way that $\mu_{\text{ess}}\tau$ is a function of only $s\tau$ and $r\tau$. This scaling rule helps us in presenting the parameter dependence of μ_{ess} in an economical way. Moreover, once a scaled result is obtained by calculations for parameters which are within computer's ability, it can be used to predict results for extreme parameters beyond computer's ability.

Figure 2 shows that this scaling rule actually holds with our model. For a number of pairs of scaled parameters $s\tau$ and $r\tau$, which were taken from a region $10^{-2} \leq s\tau \leq 10^2$ and $0 \leq r\tau \leq 10^2$, we calculated μ_{ess} for three different durations $\tau = 100, 320$, and 1000. For a fixed value of $r\tau$, the scaled results $\mu_{\text{ess}}\tau$ as a function of scaled parameter $s\tau$ lie reasonably well on a curve notwithstanding different τ 's were used.

The curve for $r\tau = 0$ (complete linkage) corresponds exactly to the contour line of evolution rate $v = \mu$ of a continuous time replicon model under a periodic selection corresponding to our (7) (Figure 3 of ISHII, MATSUDA and OGITA 1982). This is as it should be. According to the extended Haldane-Muller principle of mutation load $\partial \bar{m} / \partial \mu = 1 - v/\mu$ (ISHII and MATSUDA 1985), $v/\mu = 1$ corresponds to μ_{op} which maximizes the long-term average \bar{m} of averaged Malthusian parameter (*i.e.*, $\partial \bar{m} / \partial \mu = 0$).

Figure 2 shows that for a fixed $r\tau$, $\mu_{\text{ess}}\tau$ does not depend on $s\tau$ as long as $s\tau \leq 1$. As we further increase $s\tau$, $\mu_{\text{ess}}\tau$ decreases to a certain minimum less than 1, and then increases to 1 asymptotically in the limit of large $s\tau$. However, for $r\tau \leq 1$, the asymptotical approach of $\mu_{\text{ess}}\tau$ to 1 from below is not so conspicuous as for $r\tau \gg 1$.

The fact that $\mu_{\text{ess}} = 1/\tau$ for very strong selection has been known for completely linked modifier case $r = 0$ by LEIGH (1970). Figure 2 shows that it applies also for unlinked modifiers although the required $s\tau$ is the greater for the larger $r\tau$. For very strong selection $s \approx 1$, it can be further shown by our formula (4) that μ_{ess} is always equal to the reciprocal of the average duration of environment, whether the fluctuation is periodic or random (see APPENDIX C).

The level of $\mu_{\text{ess}}\tau$ for small $s\tau$ can be calculated analytically by assuming that for such weak selection the frequency of A allele at the main locus fluctuates only near around 0.5. By a linear analysis based on expansion of A frequencies around 0.5, we obtain in APPENDIX B.

$$\frac{1}{1 + e^{-R-2M}} \left(\frac{1 - e^{-R-2M}}{R + 2M} - 2 \frac{1 - e^{-R}}{R} \frac{e^{-2M}}{1 + e^{-2M}} \right) + \frac{1}{2M} \frac{1 - e^{-2M}}{1 + e^{-2M}} - \frac{1}{2} = 0$$

as the equation to determine $M = \mu_{\text{ess}}\tau$ as a function of $R = r\tau$ for large τ . This gives $\mu_{\text{ess}}\tau \approx 1.6061$ for $r\tau \ll 1$, and $\mu_{\text{ess}}\tau \approx 6/r\tau$ for $r\tau \gg 1$. The result for $s\tau \leq$

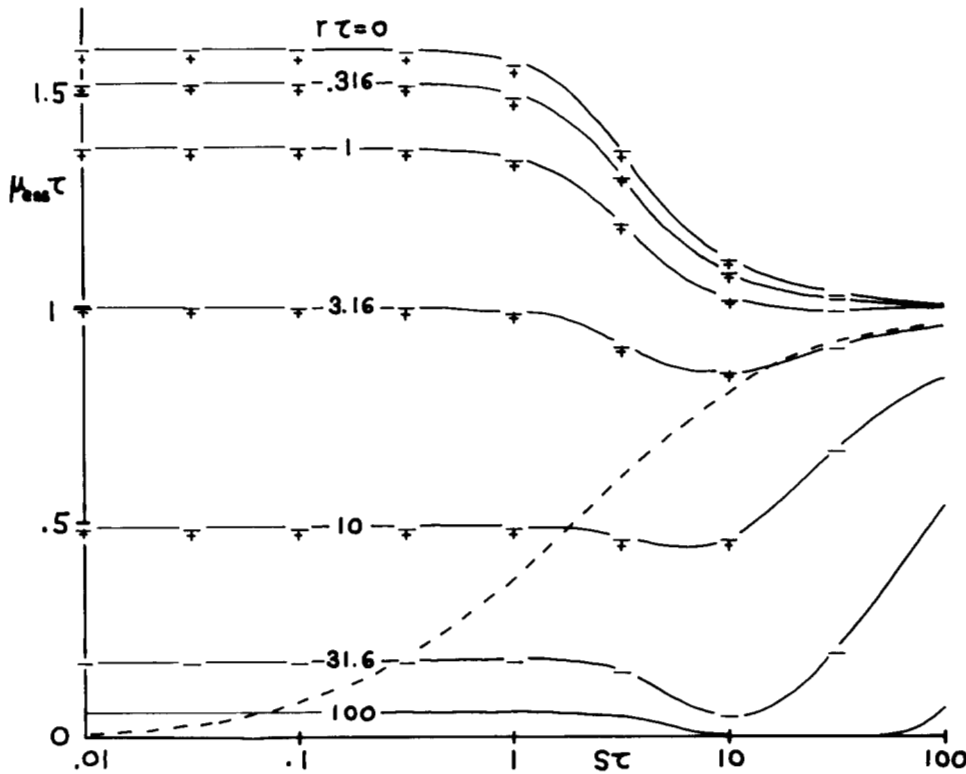


FIGURE 2.—Scaled ESS mutation rate $\mu_{\text{ess}}\tau$ vs. scaled selection strength $s\tau$ for different scaled recombination rate $r\tau = 0, 10^{-0.5}, 1, 10^{0.5}, 10, 10^{1.5}, 100$ under a periodic selection with a single strong Fourier component as in Figure 1. For a given pair of $s\tau$ and $r\tau$, μ_{ess} is calculated for three different durations $\tau = 100, 320$ and 1000 . Symbol + gives $\mu_{\text{ess}}\tau$ for $\tau = 100$, symbol - for $\tau = 320$. Lines are drawn for each fixed $r\tau$ as to connect the results for $\tau = 1000$. The broken line gives the optimal mutation rate μ_{op} as $\mu_{\text{op}}\tau$ for a random selection with strength s and average duration τ (see DISCUSSION).

1 in Figure 2 agrees very well with this analytical result.

Figure 2 also shows that $\mu_{\text{ess}}\tau$ for a fixed $s\tau$ monotonically decreases as $r\tau$ increases from 0. However, the decrease is not large as long as $r\tau \leq 1$, and becomes appreciable only for $r\tau > 1$.

The fluctuating selection (7) can be said to be *with a single strong Fourier component* (i.e., sine wave component) since its frequency spectrum has a single strong peak at the frequency $\omega = \pi/\tau$. However, there remains a question if the above result obtained for this specific case really gives us a good general picture of μ_{ess} in a fluctuating environment with a single strong Fourier component. Although we can not be very conclusive, the answer seems to be yes. For example, we may consider a purely sine wave fluctuation

$$s(t) = s \cos(\pi t/\tau + \varphi) \quad (8)$$

as a second environment. Here, s denotes the strength of selection, τ is the number of generations during which one allele is persistently favored than the other, and φ is the phase parameter of the environment. As we show in APPENDIX B, in the weak selection limit this environment gives $\mu_{\text{ess}}\tau$ which is nearly equal to $\pi/2$ for $r\tau \ll 1$, and $\pi^2/2r\tau$ for $r\tau \gg 1$. This is essentially the same result as explained in the above for the corresponding environment (7). Numerical solution of (4) for this environment shows that the behavior of $\mu_{\text{ess}}\tau$ for $s\tau > 1$ is also similar to that of the corresponding environment (7).

Fluctuation with multiple Fourier components: For a fluctuation consisting of multiple Fourier com-

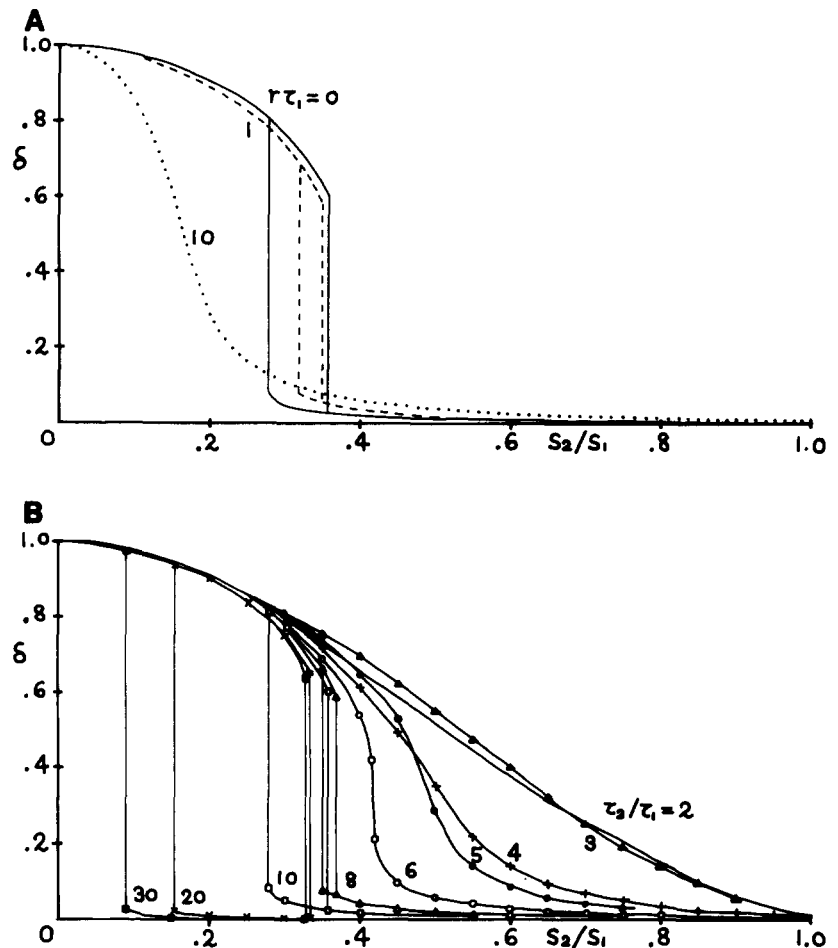
ponents, we are interested in how μ_{ess} depends on each component. As a simplest example, let us consider a two component case where $s(t)$ is the sum of two periodic sequences as are given in (7) respectively with strength and duration s_1, τ_1 and s_2, τ_2 . Denoting by $\mu_{\text{ess}}^{(i)}$ the ESS mutation rate for a periodic selection with a single strong Fourier component of strength s_i and duration τ_i , we introduce the relative deviation $\delta = (\mu_{\text{ess}} - \mu_{\text{ess}}^{(2)})/(\mu_{\text{ess}}^{(1)} - \mu_{\text{ess}}^{(2)})$ of μ_{ess} from $\mu_{\text{ess}}^{(2)}$. $\delta \approx 1$ and $\delta \approx 0$ mean that the most effective component in determining μ_{ess} is the first and the second one, respectively.

Figure 3A shows how δ changes as s_2/s_1 increases from 0 to 1. The first component is fixed as $s_1\tau_1 = 10^{-3}$ with $\tau_1 = 31$. The second component is chosen to have a fixed duration $\tau_2 = 310$. Calculations were done for $r\tau_1 = 0, 1$ and 10 . It should be noticed that the strength of the second component changes from $s_2\tau_2 = 0$ to 10^{-2} as s_2/s_1 increases from 0 to 1, and three levels of linkage correspond to $r\tau_2 = 0, 10, 100$. Since $s_1\tau_1 \ll 1$ and $s_2\tau_2 \ll 1$, $\mu_{\text{ess}}^{(1)}$ and $\mu_{\text{ess}}^{(2)}$ are almost independent of the selection strength and are essentially determined by $r\tau_1$ and $r\tau_2$, respectively.

As s_2/s_1 increases from 0 to 1, μ_{ess} always decreases monotonically, whether the modifier is linked or not. It starts from $\mu_{\text{ess}}^{(1)}$ ($\delta = 1$) and lies above $\mu_{\text{ess}}^{(2)}$ ($\delta > 0$). The mode of decrease, however, seems to make a change from sharp transition to gradual decrease as the modifier becomes more loosely linked with the main locus.

Let us examine the δ curve for $r = 0$ in more detail. In this case, $\mu_{\text{ess}}^{(1)}\tau_1 = 1.53$ and $\mu_{\text{ess}}^{(2)}\tau_2 = 1.60$. For 0.281

FIGURE 3.—A, The ESS mutation rate μ_{ess} vs. the relative strength s_2/s_1 of selection for a periodic selection with two strong Fourier components. The shorter duration component is fixed as $s_1\tau_1 = 10^{-3}$ with $\tau_1 = 31$. The longer duration component has a fixed duration $\tau_2 = 310$. The ordinate is the relative deviation $\delta = (\mu_{\text{ess}} - \mu_{\text{ess}}^{(2)}) / (\mu_{\text{ess}}^{(1)} - \mu_{\text{ess}}^{(2)})$ of μ_{ess} from $\mu_{\text{ess}}^{(2)}$, where $\mu_{\text{ess}}^{(i)}$ is the ESS mutation rate for a periodic selection with a single strong Fourier component of strength s_i and duration τ_i . The solid line is for the recombination rate $r = 0$ between the modifier and the main loci, the broken line for $r\tau_1 = 1$, and the dotted line for $r\tau_1 = 10$. B, The low pass filter effect on μ_{ess} for a periodic selection with two strong Fourier components vs. the relative duration $\tau_2/\tau_1 = 2, 3, 4, 5, 6, 8, 10, 20, 30$. Calculation is for the case of completely linked modifiers ($r = 0$). The shorter duration component is fixed in the same way as in A. The ordinate and the abscissa are taken in the same way as in A.



$\leq s_2/s_1 \leq 0.358$, it consists of two branches giving two locally stable μ_{ess} 's. At $s_2/s_1 = 0.281$ a sharp drop of δ occurs to start its lower branch, and at $s_2/s_1 = 0.358$ another sharp drop occurs to end its upper branch. Since the lower branch of δ lies below $\delta = 0.05$, we may say that the longer duration component is mainly responsible in determining μ_{ess} if s_2/s_1 exceeds 0.358. Thus, our result for the complete linkage case is in accordance with what SASAKI and IWASA (1987) found and named as the low pass filter effect as for the optimal recombination rate in a fluctuating environment with multiple Fourier components.

The low pass filter effect is observed also for unlinked modifiers. For $r\tau_1 = 1$, $\mu_{\text{ess}}^{(1)}\tau_1 = 1.32$ and $\mu_{\text{ess}}^{(2)}\tau_2 = 0.483$. Recombination does not reduce $\mu_{\text{ess}}^{(1)}$ appreciably, but reduces $\mu_{\text{ess}}^{(2)}$ significantly since $r\tau_2 = 10 \gg 1$. Even in this case, we see that δ is less than 0.05 for $s_2/s_1 \geq 0.35$. For $r\tau_1 = 10$, $\mu_{\text{ess}}^{(1)}\tau_1 = 0.402$ and $\mu_{\text{ess}}^{(2)}\tau_2 = 0.049$. Recombination reduces both $\mu_{\text{ess}}^{(1)}$ and $\mu_{\text{ess}}^{(2)}$ significantly. In this case, δ is less than 0.05 for $s_2/s_1 \geq 0.43$. However, the transition into the regime where the longer duration component dominates in determining μ_{ess} is not so sharp as for $r\tau_1 \leq 1$ cases.

Figure 3B shows δ for $r = 0$ case versus the relative strength s_2/s_1 for different $\tau_2/\tau_1 = 2, 3, 4, 5, 6, 8, 10, 20, 30$. The first component is fixed just in the same

way as in Figure 3A. δ is less than 0.05 for $s_2/s_1 \geq 0.5$ if and only if $\tau_2/\tau_1 > 5$. For a given τ_2/τ_1 , δ decreases only gradually with s_2/s_1 for $\tau_2/\tau_1 \leq 5$, but it makes sharp drops for $\tau_2/\tau_1 > 5$. The first drop of δ which starts its lower branch occurs at the smaller s_2/s_1 value for the greater τ_2/τ_1 . However, the second drop of δ which ends its higher branch occurs at an s_2/s_1 value (between 0.25 and 0.4) which does not change very much with τ_2/τ_1 .

Combining the results in Figure 3, A and B, together, we can conclude that the low pass filter effect works generally for $\tau_2/\tau_1 > 5$ and for $s_2/s_1 \geq 0.5$, whether the modifier is linked or not.

DISCUSSION

In this paper we examined the evolution of mutation rate controlled by neutral modifiers by a formula (4) for the evolutionarily stable mutation rate μ_{ess} . The main results obtained in previous sections are summarized as:

I. For any fluctuation, periodic or random: (1) if the recombination rate r per generation between the modifier and the main locus is 0, μ_{ess} is the same as the optimal mutation rate μ_{op} which maximizes the long-term geometric average of population fitness;

and (2) for any r , if the strength s of selection per generation is very large, μ_{ess} is equal to the reciprocal of the average number τ of generations (duration time) during which one allele is persistently favored than the other.

II. For a periodic fluctuation in the limit of small s and r , $\mu_{\text{ess}}\tau$ is a function of $s\tau$ and $r\tau$ with properties: (1) for a given $s\tau$, $\mu_{\text{ess}}\tau$ decreases with increasing $r\tau$; (2) for $s\tau \leq 1$, $\mu_{\text{ess}}\tau$ is almost independent of $s\tau$, and depends on $r\tau$ as $\mu_{\text{ess}}\tau \approx 1.6$ for $r\tau \ll 1$ and $\mu_{\text{ess}}\tau \approx 6/r\tau$ for $r\tau \gg 1$; and (3) for $s\tau \geq 1$, and for a given $r\tau$, $\mu_{\text{ess}}\tau$ decreases with increasing $s\tau$ to a certain minimum less than 1, and then increases to 1 asymptotically in the limit of large $s\tau$.

III. For a fluctuation consisting of multiple Fourier components (*i.e.*, sine wave components), the component with the longest period is the most effective in determining μ_{ess} (*low pass filter effect*).

Our results should be compared with those of LEIGH (1970). Based on an analysis which is valid for very strong selection, he claimed that in an asexual population selection adjusts the mutation rate toward a nonzero level which is equal to the reciprocal of the duration of fluctuating environment. Based on the general ineffectiveness of intergroup selection, he further argued that in a sexual population selection on mutation rates would operate toward zero. Our result (12) is in accordance with LEIGH's claim about an asexual population, and (II3) extends it to the general case of $s \gg 1/\tau$. However, our result does not support LEIGH's claim about a sexual population, since μ_{ess} is always positive in a periodic environment for any value of recombination rate r .

Invadability of an ESS modifier: In this paper we analyzed the evolutionary stability of a wild-type modifier against an invading modifier only at its initial stage of invasion. However, even if the invading modifier is successful when its frequency is low, it may not be so when its frequency becomes higher. In that case, the wild type modifier which is not ESS in the sense of our present analysis can persist to exist by somehow controlling the invading modifier at a low frequency level in the population. From this point, an interesting question is whether an ESS modifier with $\mu_{\text{ess}} \neq \mu_{\text{op}}$ can invade into a population of μ_{op} with the greatest average population fitness.

We studied this by a computer simulation of the whole process of modifier competition for typical cases. We found that an ESS modifier always succeeded in invading into the population of a non-ESS modifier (chosen from the attractor of the studied μ_{ess} , when more than one local ESS modifiers existed). This gives a support to our expectation that the mutation rate will evolve toward μ_{ess} as is calculated in this paper.

Random fluctuation: In the real environment some

stochastic elements are usually included in its fluctuation. So, let us consider what kind of differences will be expected in a randomly changing environment. In this case an explicit analysis of formula (4) for μ_{ess} is not so easy because $\lambda(\mu', \mu)$ in a random environment is difficult to evaluate explicitly. However, our results in (I) apply generally whether the environment is random or periodic.

Thus, if the modifier is completely linked, μ_{ess} is equal to μ_{op} . For a simple model of random environment in which the selection coefficient $s(t)$ makes a Markov process which takes two values $+s$ and $-s$ with an average duration τ , μ_{op} can be obtained directly by maximizing the long-term geometric average of population fitness which we can evaluate numerically based on an explicit result on the stationary distribution of allele A at the main locus (MATSUDA and ISHII 1981). The result $\mu_{\text{op}}\tau$ is shown in Figure 2 by a broken line. Comparing two curves of μ_{op} for a random environment and for a corresponding periodic one, we see that they are approximately at the same level for $s\tau \gg 1$. This is as is expected from our result (12). However, for weak selection $s\tau \ll 1$, we see that μ_{op} in a Markov environment approaches zero as $\mu_{\text{op}} = 0.5s$ while μ_{op} in a periodic environment stays at a finite level of $1.6061/\tau$. The smaller μ_{op} for a random environment than for a periodic one may be explained by the low pass filter effect as due to the longer period Fourier components of fluctuation which are contained in a random environment.

For the loosely linked modifier case, we can show, in a similar way as in APPENDIX B, that μ_{ess} in the above mentioned Markov environment is at most less than $10s$ in the weak selection limit ($s\tau \ll 1$). Further, since we have low pass filter effect also for unlinked modifiers, we may expect that μ_{ess} for a random environment is generally smaller than that for a corresponding periodic environment, and that the difference will be the greater for the weaker selection. This expectation was borne out by computer simulations of the modifier competition.

Using a diffusion approximation GILLESPIE (1981) studied the evolution of mutation rate by a modifier model with the selected locus under a fluctuating selection which generally brings about a marginal overdominance. Our model in the above mentioned Markov environment corresponds in the weak selection limit $s\tau \ll 1$ to his diffusion model with parameters $A = 0$ and $B = 1$ (MATSUDA and ISHII 1981), for which his result is that selection will continue to increase the mutation rate whatever its current value. This result is at variance with our above result of $\mu_{\text{ess}} < 10s$. Since the diffusion approximation used by him is justifiable only for $\mu/s^2\tau = O(1)$ in the limit of $s \rightarrow 0$ and $s\tau \rightarrow 0$, his result is of dubious significance for

μ such as $\mu/s = O(1)$. We presume that this is the cause of the discrepancy.

Cost of preventing mutation: In this paper we have assumed that mutation rate modifiers are selectively neutral, but they can not be always neutral for real organisms. In order to reduce the mutation rate, it may be necessary for organisms to develop a replication system where the replication error is reduced. This necessarily requires more time and free energy for replication, causing a decrease in the multiplication rate of organisms per unit time. Then, μ_{ess} established by such non-neutral modifiers is expected not to be so low as by neutral ones.

As a simple model of non-neutral modifiers, let us consider a multiplicative two-locus model with a modifier of mutation rate μ contributing a fitness component $f(\mu)$. Then, the formula (4a) is modified as

$$\Delta(\mu)/\lambda(\mu, \mu) + c(\mu) = 0. \quad (4a')$$

Here, $c(\mu) \equiv f'(\mu)/f(\mu)$ is the relative cost to reduce mutation rate by a unit amount, and can be assumed to be non-negative. Then, for $c(\mu) > 0$ formula (4a') gives μ_{ess} larger than in the case of neutral modifiers ($c(\mu) = 0$) under the same selection pressure acting at the main locus.

If the modifier is loosely linked with the main locus, we found that μ_{ess} by neutral modifiers is smaller than the optimal rate μ_{op} . Then, we may ask how much cost c^* is needed to increase μ_{ess} to the level of μ_{op} of the neutral modifier case. According to (4a'), c^* is given by $-\Delta(\mu_{\text{op}})/\lambda(\mu_{\text{op}}, \mu_{\text{op}})$.

Figure 4 shows for the periodic selection (7) how the critical cost c^* depends on the scaled model parameters $s\tau$ and $r\tau$. For a given $s\tau$, c^* increases with increasing $r\tau$ from $c^* = 0$ at $r\tau = 0$ to a certain maximum \bar{c}^* , and then decreases to 0 asymptotically in the limit of large $r\tau$. Curves for different $s\tau$ values less than 1 are of the same shape but are proportional to $(s\tau)^2$. The maximum \bar{c}^* for a given $s\tau$ depends on $s\tau \leq 1$ as $\bar{c}^* \approx (s\tau)^2/100$. For $1 \leq s\tau \leq 10^6$, \bar{c}^* gradually increases with increasing $s\tau$ but stays less than 1. The behavior of c^* for $s\tau \leq 1$ can be explained by the result of APPENDIX B for the weak selection limit.

Based on an extended Haldane-Muller principle of mutation load, ISHII and MATSUDA (1985) proved that, if there is no cost of preventing mutation, μ_{op} is equal to the evolution rate v which is defined as the increase rate of population average of mutation numbers which have occurred along the phylogenetic line leading to the present replicons. They further showed that $v < \mu_{\text{op}}$ if there is a positive cost of preventing mutation, and pointed out that this may explain, from a selectionist perspective, the fact that molecular evolution rates are smaller than total mutation rates. Their argument assumed completely linked mutation rate modifiers, but can be extended to the case of unlinked modifiers as $v \geq \mu_{\text{ess}}$ for $c^* \geq c$. Thus, the

above mentioned fact about molecular evolution rates corresponds to the case when the unlinked modifiers incur a positive cost c of preventing mutation greater than c^* . When some reliable data are obtained in the future about how large is the cost c , it can be compared with the values of c^* given in Figure 4.

Two allele model under fluctuating selection: Our model assumes at the main locus that mutation occurs between two alleles, and that they become the more fit than the other alternatively. This assumption is very suitable for the *flip-flop* mutation in bacteria and bacteriophage (WATSON *et al.* 1987).

For example, individual *Salmonella* bacteria can alternate flagella protein expression between two types, H1 and H2, which differ in antigen property. Since the flagella protein is a dominant antigen of *Salmonella*, the switching is favorable in eluding the host immune defense. The molecular mechanism has been clarified (BORST and GREAVES 1987): In one phase of gene expression, the gene for H1 is transcribed together with an adjacent gene that codes for the repressor of gene for H2—hence only H1 is expressed. In alternative phase, the promoter sequence of H1 is inverted, and neither H1 gene nor repressor gene for H2 is transcribed, then only H2 gene is expressed. The orientation of the invertible segment thus determines gene expression of flagella protein. The rate of occasional inversion of the segment is regulated by the recombinase, *Hin*, which is coded within the segment itself, together with promoter sequence of H1. Hence, this is an example of completely linked mutation rate modifier in our model.

Another mechanism for switching between two alleles is a *cassette mechanism* for the mating type of yeast (DARNELL 1982). It is known that there are two silent loci together with a single expression locus. The expression locus is occasionally renewed by gene conversion from one of the two silent loci in which the information is stored.

If the rate of flip-flop in bacteria or mating type change in yeast is to be evolutionarily determined, our model gives the rate which is to be realized as a result of evolution.

General interpretation of the two-allele model—parity model: At first sight, or taken literally, the two-allele model for the main locus as studied in this paper may sound too simple or too artificial to get an insight into the real mutation process occurring throughout the genome of organisms. One may mention the possibility that most mutations will be deleterious in all or essentially all environments, contrary to our assumption of fluctuating environments. One may argue then that even if there are a few sites or loci which behave something like the main locus of the model considered, the applicability of the model to the prob-

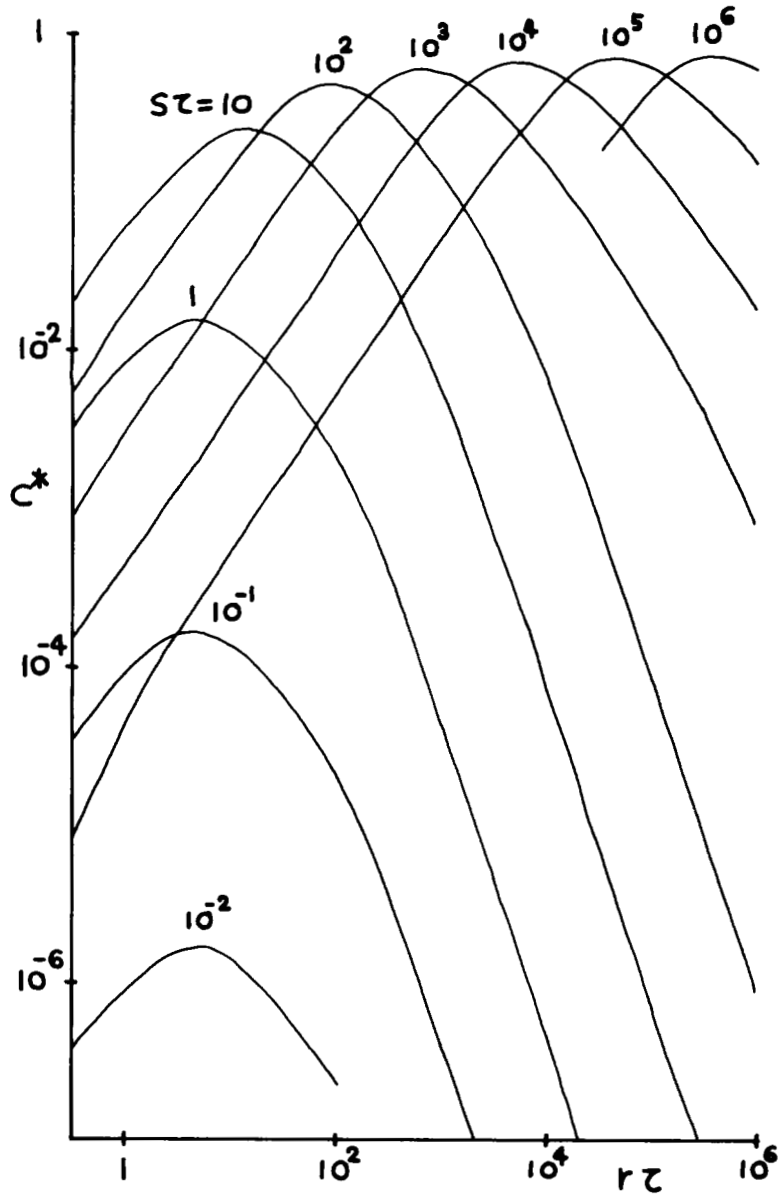


FIGURE 4.—The critical cost c^* of preventing mutation that is needed to increase μ_{ess} to the level of μ_{op} of the neutral modifier case. The same periodic selection with a single strong Fourier component as in Figure 1 is assumed. The scaled result for c^* is calculated for a given pair of scaled parameters $s\tau$ and $r\tau$ with $\tau = \max(10s\tau, 10r\tau)$. Lines are drawn for each fixed $s\tau$ value.

lem of mutation rate may be very limited unless the modifier locus is extremely site specific in its actions.

We do not deny the possibility that many mutations will be deleterious in essentially all environments, yet we consider that models studied in this paper may still represent some essential feature of the real replication unit (*replicon*) such as a chromosome or DNA by the following reasons.

We have assumed only two alleles at the main locus, but their fitnesses are generally time-dependent. Then, by classifying all the possible genetic states of replicons into just two types, we can regard the above fitness as an average fitness of a subpopulation consisting of the respective type of replicons. In that case, the fitness of replicons which are deleterious in every environment will be considerably low, so that their frequencies will remain very small in each subpopu-

lation. Then, their effect on the average fitness of each subpopulation will also be small.

Therefore, under a suitable dichotomous classification of genetic states, it may be possible that the average fitness of each type fluctuates essentially like the model we have considered in this paper. If this is indeed the case, the mutation rate modifier need not be site specific in order for our model to be applicable.

For instance, we may classify four kinds of bases A, T, G, and C into purines (A or G) and pyrimidines (T or C). Then, we may classify the base sequences of DNA according to whether the total number of purines contained in each sequence is even or odd. It is like the classification of the internal states of elementary particles by even and odd parity.

In order to get a simple and unified view of molecular evolution from the population genetical standpoint, we consider that further study of such simple

models with general interpretation as may be called a "parity model" will be worthwhile.

We thank S. KASHIWAMURA and J. ISHIMOTO for helpful comments on an earlier version of this paper. We also thank anonymous reviewers of this paper for helpful comments to improve the presentation. This work was supported in part by a Grant-in-Aid for Special Project Research on Biological Aspects of Optimal Strategy and Social Structure from Ministry of Education, Science and Culture.

LITERATURE CITED

- BORST, P., and D. R. GREAVES, 1987 Programmed gene rearrangements altering gene expression. *Science* **235**: 658–667.
- DARNELL, J. E., JR., 1982 Variety in the level of gene control in eucaryotic cells. *Nature* **297**: 365–370.
- GILLESPIE, J., 1981 Mutation modification in a random environment. *Evolution* **35**: 468–476.
- ISHII, K., and H. MATSUDA, 1985 Extension of the Haldane-Muller principle of mutation load with application for estimating a possible range of relative evolution rate. *Genet. Res.* **46**: 75–84.
- ISHII, K., H. MATSUDA and N. OGITA, 1982 A mathematical model of biological evolution. *J. Math. Biol.* **14**: 327–353.
- KIMURA, M., 1960 Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load. *J. Genet.* **57**: 21–34.
- KIMURA, M., 1967 On the evolutionary adjustment of spontaneous mutation rates. *Genet. Res.* **9**: 23–34.
- LEIGH, E. G., 1970 Natural selection and mutability. *Am. Nat.* **104**: 301–305.
- LEIGH, E. G., 1973 The evolution of mutation rates. *Genetics* **73** (Suppl.): s1–s18.
- LEVINS, R., 1967 Theory of fitness in a heterogeneous environment. VI. The adaptive significance of mutation. *Genetics* **56**: 163–178.
- LIEBERMAN, U., and M. W. FELDMAN, 1986 Modifiers of mutation rate: a general reduction principle. *Theor. Popul. Biol.* **30**: 125–142.
- MATSUDA, H., and K. ISHII, 1981 Stationary gene frequency distribution in the environment fluctuating between two distinct states. *J. Math. Biol.* **11**: 119–141.
- MAYNARD SMITH, J., and G. R. PRICE, 1973 The logic of animal conflict. *Nature* **246**: 15–18.
- SASAKI, A., and Y. IWASA, 1987 Optimal recombination rate in fluctuating environments. *Genetics* **115**: 377–388.
- STURTEVANT, A. H., 1937 Essays on evolution. I. On the effects of selection on mutation rate. *Q. Rev. Biol.* **12**: 464–467.
- TAKAHATA, N., K. ISHII, and H. MATSUDA, 1975 Effect of temporal fluctuation of selection coefficient on gene frequency in a population. *Proc. Natl. Acad. Sci. USA* **72**: 4541–4545.
- WATSON, J. D., A. H. HOPKINS, J. W. ROBERTS, J. A. STEITZ and A. M. WEINER, 1987 *Molecular Biology of the Gene*, Vol. 1, Ed. 4. Benjamin/Cummings, Menlo Park, Calif.

APPENDIX A

Method of numerical analysis of (4)

First, let us write down the characteristic equation of M_T as

$$\lambda^2 - \lambda \text{Tr}M_T + \det M_T = 0. \quad (\text{A1})$$

Its greatest eigenvalue $\lambda(\mu', \mu)$ is given by

$$\lambda(\mu', \mu) = \{\text{Tr}M_T + \sqrt{(\text{Tr}M_T)^2 - 4\det M_T}\}/2. \quad (\text{A2})$$

Then, differentiating (A1) with μ' , we have

$$(2\lambda - \text{Tr}M_T)\partial\lambda/\partial\mu' - \lambda\partial\text{Tr}M_T/\partial\mu' + \partial\det M_T/\partial\mu' = 0,$$

which gives

$$\partial\lambda(\mu', \mu)/\partial\mu' = (\lambda\partial\text{Tr}M_T/\partial\mu' - \partial\det M_T/\partial\mu')/(2\lambda - \text{Tr}M_T). \quad (\text{A3})$$

Thus, in order to calculate by (A2) and (A3) the value of $\Lambda(\mu) = [\partial\lambda(\mu', \mu)/\partial\mu']_{\mu'=\mu}$ for a given μ , we need only to know four values at $\mu' = \mu$ of $\text{Tr}M_T$ and $\det M_T$ together with their partial derivatives with respect to μ' . $\text{Tr}M_T$ and $\partial\text{Tr}M_T/\partial\mu'$ are immediately obtained once matrices M_T and $\partial M_T/\partial\mu'$ are numerically calculated by the following iteration

$$M_{t+1} = M(t)M_t, \quad (\text{A4a})$$

and

$$\partial M_{t+1}/\partial\mu' = M(t)\partial M_t/\partial\mu' + [\partial M(t)/\partial\mu']M_t \quad (\text{A4b})$$

for $t = 0, 1, \dots, T-1$, where $\partial M(t)/\partial\mu'$ is calculated from (2) as

$$\partial M(t)/\partial\mu' = (1-r) \begin{bmatrix} -1-s(t) & 1-s(t) \\ 1+s(t) & -1+s(t) \end{bmatrix}. \quad (\text{A5})$$

$\det M_T$ is calculated from (2) and (6) as

$$\det M_T = \{(1-r)(1-2\mu')\}^{2r} \prod_{t=0}^{T-1} \{1-s(t)\}^2, \quad (\text{A6a})$$

with

$$\partial\det M_T/\partial\mu' = -4r\det M_T/(1-2\mu'). \quad (\text{A6b})$$

Calculating necessary μ' derivatives by (A4)–(A6), we obtain by (A2) and (A3) the value of $\Lambda(\mu)$ for each μ . Finally, zeros of $\Lambda(\mu)$ in the range $[0, 1]$ are obtained by a routine of bisection.

APPENDIX B

μ_{ess} in the weak selection limit

In this appendix we show how we can analytically study μ_{ess} when the frequencies $x(t)$ and $y(t)$ of A alleles, respectively, among B carrying genomes and among b carrying ones, keep fluctuating only very near around 0.5. Such a situation is realized when the strength s of selection is very small compared with the mutation rate μ .

First, we restate the condition (4) for μ_{ess} in terms of $x(t)$ and $y(t)$. Since the average fitness $\bar{w}(t)$ of b carrying genomes is expressed as $\bar{w}(t) = 1 + s(t)\{2y(t) - 1\}$ in terms of $y(t) = N_1(t)/N(t)$, the increase rate $\lambda(\mu', \mu)$ as the long-term geometric average of $\bar{w}(t)$ is given in the weak selection limit as

$$\lambda(\mu', \mu) = 1 + \lim_{t \rightarrow \infty} t^{-1} \sum_{t'=0}^{t-1} s(t')\{2y(t') - 1\},$$

where we have neglected higher order terms of s . Compar-

ing this with the rate $\lambda(\mu, \mu')$ for B carrying genomes, which is obtained by replacing $y(t')$ with $x(t' - 1)$ in the above, we find

$$\langle (y - x)s \rangle \equiv \lim_{t \rightarrow \infty} t^{-1} \sum_{t'=0}^{t-1} \{y(t') - x(t' - 1)\} s(t') < 0 \quad (\text{B1})$$

for $\mu' \neq \mu$

as the condition for μ to be a μ_{ess} .

From (1) and (2), the time change of $y(t)$ is given by

$$y(t + 1) = [rx(t) + (1 - r)(1 - \mu')]\{1 + s(t)\}y(t) + \{rx(t) + (1 - r)\mu'\}\{1 - s(t)\}\{1 - y(t)\} / [1 + s(t)\{2y(t) - 1\}].$$

The time change of $x(t)$ is given by (3) since $x(t) = x_1(t)$. By assumption, the deviations of $x(t)$ and $y(t)$ from 0.5 become small magnitude quantities after a finite transient period. Therefore, by expanding $x(t)$ and $y(t)$ around 0.5 and keeping only the leading terms in the equations for the time change of $x(t)$ and $y(t)$, we find that $\xi(t) = x(t - 1) - 0.5$ and $\zeta(t) = y(t) - x(t - 1)$ become to satisfy

$$\xi(t + 1) = (1 - 2\mu)\{s(t)/2 + \xi(t)\} \quad (\text{B2})$$

and

$$\zeta(t + 1) = (1 - r)[-(\mu' - \mu)\{s(t) + 2\xi(t)\} + (1 - 2\mu')\zeta(t)] \quad (\text{B3})$$

to the lowest order terms of s .

Alternating environment with duration τ : We consider a periodic selection with $s(t)$ given by (7). From the assumed symmetry of $s(t)$, $\xi(t)$ and $\zeta(t)$ converge to the periodic sequences which satisfy $\xi(t + \tau) = -\xi(t)$ and $\zeta(t + \tau) = -\zeta(t)$ for any t . Then, the average $\langle \zeta s \rangle$ in (B1) is calculated as

$$\langle \zeta s \rangle = s\tau^{-1} \sum_{t=0}^{\tau-1} \zeta(t) \quad (\text{B4})$$

in terms of a sequence $\{\zeta(t); t = 0, 1, \dots, \tau\}$ which is the solution of constant coefficient equations

$$\zeta(t + 1) = (1 - 2\mu)\{s/2 + \xi(t)\} \quad (t = 0, 1, \dots, \tau - 1) \quad (\text{B5})$$

and

$$\zeta(t + 1) = (1 - r)[-(\mu' - \mu)\{s + 2\xi(t)\} + (1 - 2\mu')\zeta(t)] \quad (t = 0, 1, \dots, \tau - 1) \quad (\text{B6})$$

satisfying boundary conditions $\xi(\tau) = -\xi(0)$ and $\zeta(\tau) = -\zeta(0)$.

From (B5) we find $\xi(t)$ as

$$\xi(t) = (1 - 2\mu)^t \xi(0) + s(1 - 2\mu)\{1 - (1 - 2\mu)^t\} / 4\mu, \quad (\text{B7a})$$

where $\xi(0)$ is determined as

$$\xi(0) = -s(1 - 2\mu)\{1 - (1 - 2\mu)^\tau\} / 4\mu\{1 + (1 - 2\mu)^\tau\} \quad (\text{B7b})$$

from the boundary condition. Then we substitute (B7) in (B6) and find $\zeta(t)$ in the limit of $\mu' \rightarrow \mu$ as

$$\zeta(t) = \{(1 - r)(1 - 2\mu)^t\} \zeta(0) + (1 - r)s \frac{\mu' - \mu}{\mu} \cdot \left[\frac{1 - (1 - r)^t}{r} \frac{(1 - 2\mu)^t}{1 + (1 - 2\mu)^t} - \frac{1}{2} \frac{1 - \{(1 - r)(1 - 2\mu)\}^t}{1 - (1 - r)(1 - 2\mu)} \right], \quad (\text{B8a})$$

where $\zeta(0)$ is determined as

$$\zeta(0) = - \frac{(1 - r)s(\mu' - \mu)/\mu}{1 + \{(1 - r)(1 - 2\mu)\}^\tau} \cdot \left[\frac{1 - (1 - r)^\tau}{r} \frac{(1 - 2\mu)^\tau}{1 + (1 - 2\mu)^\tau} - \frac{1}{2} \frac{1 - \{(1 - r)(1 - 2\mu)\}^\tau}{1 - (1 - r)(1 - 2\mu)} \right] \quad (\text{B8b})$$

from the boundary condition. Substituting (B8) in (B4), we finally obtain $\langle \zeta s \rangle$ as

$$\langle \zeta s \rangle = \frac{(1 - r)s^2}{1 - (1 - r)(1 - 2\mu)} \frac{\mu' - \mu}{\mu} f(\mu, r) \quad (\text{B9a})$$

with

$$f(\mu, r) = \frac{1/\tau}{1 + \{(1 - r)(1 - 2\mu)\}^\tau} \cdot \left[\frac{1 - \{(1 - r)(1 - 2\mu)\}^\tau}{1 - (1 - r)(1 - 2\mu)} - 2 \frac{1 - (1 - r)^\tau}{r} \frac{(1 - 2\mu)^\tau}{1 + (1 - 2\mu)^\tau} \right] + \frac{1 - 2\mu}{2\tau\mu} \frac{1 - (1 - 2\mu)^\tau}{1 + (1 - 2\mu)^\tau} - \frac{1}{2}. \quad (\text{B9b})$$

According to the result (B9a) for $\langle \zeta s \rangle$, the condition (B1) for μ to be a μ_{ess} becomes as $f(\mu, r) = 0$ for $0 < \mu < 1$, and $f(1, r) \geq 0$ for $\mu = 1$. The result (B9b) for $f(\mu, r)$ then shows that $\mu = 1$ is a μ_{ess} if and only if τ is an odd integer. We further see that $f(\mu, r) > 0$ for $0 < \mu < 1$ if $\tau = 1$, and that $f(\mu, r) = 0$ only for $\mu = 0.5$ among $0 < \mu < 1$ if $\tau = 2$. Therefore, in the weak selection limit, $\mu = 1$ is the only μ_{ess} for $\tau = 1$, and $\mu = 0.5$ is the only one for $\tau = 2$. For $\tau \geq 3$, $f(\mu, r) = 0$ has a root in $0 < \mu < 1$, giving a μ_{ess} which generally depends on two parameters τ and r . These results explain the numerical results for $s\tau \leq 1$ in Figure 1.

We next show how the last mentioned μ_{ess} becomes to satisfy a scaling rule for $\tau \gg 1$. We fix $\mu\tau = M$ and $r\tau = R$ at finite values and let $\tau \rightarrow \infty$. Then, we find

$$F(M, R) \equiv \lim_{\tau \rightarrow \infty} f(M/\tau, R/\tau) = \frac{1}{1 + e^{-R-2M}} \left(\frac{1 - e^{-R-2M}}{R + 2M} - 2 \frac{1 - e^{-R}}{R} \frac{e^{-2M}}{1 + e^{-2M}} \right) + \frac{1}{2M} \frac{1 - e^{-2M}}{1 + e^{-2M}} - \frac{1}{2}. \quad (\text{B10})$$

For a given $R = r\tau$, $F(M, R)$ as a function of M has a zero $M > 0$ which gives the scaled $\mu_{\text{ess}\tau}$. In a special case of $R = 0$, which corresponds to the completely linked modifier case, $F(M, 0) = 0$ is numerically solved to give $M = 1.6061 \dots$. In the opposite case of $R \gg 1$, which corresponds to the loosely linked modifier case, $M \ll 1$ satisfies $M/R - M^2/6 = 0$ to its second order terms, with a positive zero $M = 6/R$.

These analytical results explain very well the numerical

results for $s\tau \leq 1$ in Figure 2. There, we further note that $\mu_{\text{ess}}\tau$ for a fixed $r\tau$ is independent of $s\tau$ not only for $s < \mu$ but also for $s \geq \mu$ as long as $s\tau \leq 1$. This suggests that periodic selection for $s\tau < 1$ also can produce an additional drift of $x(t)$ toward 0.5 just as is known for stochastic selection (GILLESPIE 1972; TAKAHATA, ISHII and MATSUDA 1975).

Sinusoidally oscillating environment with period 2τ : We now consider a second example of periodic selection with $s(t)$ given by (8). This environment is similar to the previous one in that one allele is persistently favored than the other for τ generations. Noting that $\cos(\omega t + \varphi) = \text{Re}(e^{i(\omega t + \varphi)})$, we assume the deviation of frequencies as $\xi(t) = \text{Re}(sXe^{i(\omega t + \varphi)})$ and $\zeta(t) = \text{Re}(sZe^{i(\omega t + \varphi)})$ with $\omega = \pi/\tau$. Then, the average in (B1) is calculated as $\langle \xi s \rangle = s^2 \text{Re}(Z/2)$, while the complex amplitudes X and Z satisfy from (B2) and (B3)

$$Xe^{i\omega} = (1 - 2\mu)(1/2 + X) \quad (\text{B11a})$$

and

$$Ze^{i\omega} = (1 - r)\{-(\mu' - \mu)(1 + 2X) + (1 - 2\mu')Z\}. \quad (\text{B11b})$$

After some calculation we obtain in the limit of $\mu' \rightarrow \mu$

$$\begin{aligned} \langle \xi s \rangle &= (1 - r)(\mu' - \mu)s^2 g(\mu, r)/2\{1 + (1 - 2\mu)^2 \\ &\quad - 2(1 - 2\mu)\cos \omega\} \times \{1 + (1 - r)^2(1 - 2\mu)^2 \\ &\quad - 2(1 - r)(1 - 2\mu)\cos \omega\} \end{aligned} \quad (\text{B12a})$$

with

$$\begin{aligned} g(\mu, r) &= (2 - r)(1 - 2\mu) \\ &\quad - \{(1 - r)(1 - 2\mu)^2 + 1\}\cos \omega. \end{aligned} \quad (\text{B12b})$$

According to the result (B12a) for $\langle \xi s \rangle$, the condition (B1) for μ to be a μ_{ess} becomes as $g(\mu, r) = 0$ for $0 < \mu < 1$, and $g(1, r) \geq 0$ for $\mu = 1$. The result (B12b) for $g(\mu, r)$ then shows that $\mu = 1$ is a μ_{ess} if and only if $\tau = 1$. We further see that $g(\mu, r) > 0$ for $0 < \mu < 1$ if $\tau = 1$, and that $g(\mu, r) = 0$ only for $\mu = 0.5$ among $0 < \mu < 1$ if $\tau = 2$. Therefore, in the weak selection limit, $\mu = 1$ is the only μ_{ess} for $\tau = 1$, and $\mu = 0.5$ is the only one for $\tau = 2$. For $\tau \geq 3$, $g(\mu, r) = 0$ has a root in $0 < \mu < 1$, giving a μ_{ess} which generally depends on two parameters τ and r .

The scaled result for this environment is given by $G(M, R) = 0$ with

$$\begin{aligned} G(M, R) &= \lim_{\tau \rightarrow \infty} \tau^2 g(M/\tau, R/\tau) \\ &= \pi^2 - 2RM - 4M^2. \end{aligned} \quad (\text{B13})$$

Thus, we find $M = (\sqrt{R^2 + 4\pi^2} - R)/4$. This gives $M = \pi/2$ for $R \ll 2\pi$, and $M = \pi^2/2R$ for $R \gg 2\pi$. It should be

noted that the results for this environment are essentially the same as those for the corresponding alternating environment.

Continuous time models: If we are interested only in the scaled results obtained in the above by taking a limit of $\tau \rightarrow \infty$, it should be noted that they can be obtained more easily by starting directly from a continuous time model corresponding to our discrete time model (1) and (2). Then the rate λ becomes equal to the long-term arithmetic average of the average Malthusian parameter $\bar{m}(t)$ of b carrying genomes in a population dominated by B carrying ones. The linear analysis based on the expansion of frequencies around 0.5 can be easily carried out in a similar manner as has been explained in the above for discrete time models.

APPENDIX C

μ_{ess} under very strong selection

In this appendix we consider a fluctuating selection where $s(t)$ takes only two values $+s$ and $-s$ with an average duration τ . The fluctuation may be either periodic or random.

Under very strong selection with the strength of selection s very near to 1, we find that the frequency $x(t' - 1)$ of A alleles among B carrying genomes, just before recombination takes place in the $(t' - 1)$ th generation, is $1 - \mu$ or μ according to whether $s(t' - 1)$ is $+s$ or $-s$. Then, the frequency $y(t)$ of A alleles among b carrying genomes, just before selection takes place in the t th generation, is y_+ or y_- according to whether $s(t' - 1)$ is $+s$ or $-s$. Therefore, the average fitness $\bar{w}(t) = 1 + s(t)\{2y(t) - 1\}$ of b carrying genomes is $2y_+$ or $2y_-$ according to whether $s(t) = s(t - 1)$ or not.

Since the environment continues to be the same for τ generations on the average, we find that the rate $\lambda(\mu', \mu)$ as the long-term geometric average of $\bar{w}(t)$ is given by

$$\log \lambda(\mu', \mu) = \log 2 + \{\log y_- + (\tau - 1) \log y_+\}/\tau.$$

Substituting this in (4), and noting that

$$\Lambda(\mu) = (1 - r)\lambda(1 - \tau\mu)/\tau\mu(1 - \mu),$$

we obtain $\mu_{\text{ess}} = 1/\tau$ for any recombination rate r between the modifier and the main locus.

It can be shown in a similar way that the same result as above holds also for the following more general situation. We assume that there are a finite number a of alleles at the selected locus with the fitness of allele A_i given by $1 + s_i(t)$ and the mutation rate from A_i to A_j given by $\mu_j f_{ij}$, where μ is the total mutation rate specified by an allele at the modifier locus and f_{ij} is a non-negative constant satisfying $\sum_j f_{ij} = 1$. As a very strong fluctuating selection, we assume such that at a time all alleles except one is nearly lethal with $s_i(t) \approx -1$ and the nonlethal allele changes with time with an average duration τ .