THEORY OF FITNESS IN A HETEROGENEOUS ENVIRONMENT. VI. THE ADAPTIVE SIGNIFICANCE OF MUTATION[1]

RICHARD LEVINS

University of Puerto Rico, Rio Piedras[2]

Received January 9, 1967

RECENT work on mutation has stressed its harmful effects. Since under a wide range of assumptions natural selection in a constant environment results in a genetic equilibrium at optimal gene frequencies (that is, those gene frequencies which maximize the adaptive value $\overline{W}$), the only effect of mutation is to displace the gene frequencies from their optima and reduce fitness by an amount which is designated mutational load. It is also generally recognized that mutation is the ultimate source of genetic variability and hence of adaptation to new conditions. Thus there would seem to be some optimal mutation rate at which the harmful effects (load) are offset by the advantage a mutant allele may have in some new environment.

KIMURA (1960) was the first to consider simultaneously the loss of fitness due to mutational load and the contribution that mutation makes in the adaptation to new conditions. He considered secular evolutionary change in which genes are being replaced at a rate corresponding to HALDANE's (1957) estimate of one allele per 300 generations. KIMURA allowed almost complete dominance and mutation toward the favored allele only. This led him to an estimated total mutation rate over all relevant loci of about 0.06 per generation.

Our study differs from KIMURA's in that we are concerned with the contribution of mutation to the short term survival of a population under fluctuating conditions, the pattern of which persists for a long time. Therefore our results are sensitive to the pattern of environmental fluctuation, especially the variance and autocorrelation of the environment.

Consider a single locus with two alleles in a random mating population. At any given time the mean fitness of the population is

$$1.01 \qquad \overline{W} = W_{11}x^2 + 2W_{12}x(1-x) + W_{22}(1-x)^2$$

where x is the frequency of the allele A and W_{11}, W_{12} and W_{22} are the fitnesses of the genotypes AA, Aa, and aa, respectively.

If the W_{ij} are random variables without correlation from one generation to the next, the average fitness is given by

$$1.02 \quad E(\overline{W}) = \overline{W}_{11}\bar{x}^2 + 2\overline{W}_{12}\bar{x}(1-\bar{x}) + \overline{W}_{22}(1-\bar{x})^2 + (\overline{W}_{11}+\overline{W}_{22}-2\overline{W}_{12})\sigma_x^2$$

where the bars above symbols indicate expected values. In order to have a steady state distribution of gene frequencies instead of the fixation of one allele, we

[1] This work was partly supported by U.S. Atomic Energy Commission Contract AT(30-1)-2620.

[2] Present address: Department of Zoology and Committee on Mathematical Biology, University of Chicago, Chicago, Illinois 60637.

specify an average heterosis so that the coefficient of σ_x^2 is negative. Thus the fitness is reduced by fluctuation of gene frequency. If there are correlations between the W_{ij} of successive generations, there will also be correlations between the W_{ij} and x, so that their covariances will appear in equation 1.02.

We will consider two models. In model I we concentrate on the effect of mutation on the mean gene frequency. It will be shown that under fluctuating conditions the average gene frequency is not the optimum for the average environment, so that mutation can increase fitness by bringing the average closer to optimum. We demonstrate this for the most unfavorable situation, a heterotic lethal. In model II the emphasis is on the covariance of gene frequency and environment. Here we make special assumptions of symmetry so that mutation does not affect the mean gene frequency. Both models lead to optimum mutation rates which are much too high. Possible explanations for this are considered in the DISCUSSION. Finally, we examine the circumstances under which natural selection can increase the mutation rate.

Model I. Heterotic lethal

Here we set $W_{11} = 0$, $W_{12} = 1$, and $W_{22} = 1-s$. If s were constant, say $\bar{s}$, equilibrium of the lethal allele would be reached at

1.03 $$\hat{x} = \bar{s}/(1+\bar{s})$$

However, we assume that s is a random variable with mean $\bar{s}$, variance σ_s^2, and no autocorrelation. At any one time,

1.04 $$\overline{W} = 1 - x^2 - s(1-x)^2$$

This has the expected value

1.05 $$E(W) = 1 - \bar{s} + 2\bar{s}\bar{x} - (1+\bar{s})\bar{x}^2 - (1+\bar{x})\sigma_x^2$$

For overlapping generations the rate of change of the population is small despite the intense selection against the lethal, and we can use the continuous approximation for the rate of change:

1.06 $$\frac{dx}{dt} = \tfrac{1}{2}\, x(1-x)[s - (1+s)x]$$

The average gene frequency also changes. Taking the expected value of both sides of 1.06, and noting that $E(dx/dt) = d(E[x])/dt$, we have

1.07 $$\frac{d\bar{x}}{dt} = \tfrac{1}{2}\{\bar{x}(1-\bar{x})[\bar{s}-(1+\bar{s})\bar{x}] - \sigma_x^2(2\bar{s}+1) + (1+\bar{s})[3\bar{x}\sigma_x^2 + \mu_3]\}$$

where μ_3 is the third moment of x about its mean. When the mean value, $\bar{x}$, is equal to its optimum, $\hat{x} = \bar{s}/(1+\bar{s})$, this becomes

1.08 $$\frac{d\bar{x}}{dt} = \tfrac{1}{2}\,[-(1-\bar{s})\sigma_x^2 + (1+\bar{s})\mu_3]$$

For $\bar{s}$ very small (that is, for the viable homozygote nearly as fit as the heterozygote on the average), this is roughly

1.09 $$\frac{d\bar{x}}{dt} \sim \tfrac{1}{2}\,[\mu_3 - \sigma_x^2 < 0]$$

Hence $d\bar{x}/dt$ decreases from $\hat{x}$, and the steady state mean $\bar{x}$ is less than $\hat{x}$.

Since the mean lethal gene frequency is below the optimum, mutation toward

the lethal may increase fitness. In order to determine quantitative values, we used the steady state distribution as given by WRIGHT (1931):

$$1.10 \qquad \Phi(x) = \frac{1}{V} exp\,(2\int \frac{M}{V} dx)$$

where M, the instantaneous expected change in gene frequency, here is $\frac{1}{2}\, x(1-x)\,[\bar{s} - (1+\bar{s})x] + (1-x)u$ for mutation rate u toward the lethal, and V is the instantaneous variance, here $\frac{1}{4}\, x^2\,(1-x)^4\,\sigma_s^2$. Numerical computation on the IBM 7074 gave us $\bar{x}$, σ_x^2, and $E(\overline{W})$ for different values of $\bar{s}$, σ_s^2, and u (the mutation rate to the lethal). These are shown in Table 1.

TABLE 1

The effect of mutation rate on average fitness, mean gene frequency, and variance of gene frequency

s	σ_s^2	x	u	$E(W)$	$\bar{x}$	σ_x^2
.05	.03	.048	0	.9513	.0172	.000130
			.0001	.9514	.0193	.000133
			.0005	.9517	.0249	.000133
			.0010	.9519	.0289	.000145
. .	.04	.048	0	.9511	.0150	.000142
			.0001	.9512	.0169	.000155
			.0005	.9516	.0234	.000168
			.0010	.9518	.0281	.000178
			.0020	.9520	.0344	.000207
			.0030	.9521	.0397	.000227
			.0040	.9521	.0442	.000242
			.0050	.9521*	.0482	.000256
			.0060	.9521	.0518	.000269
.10	.03	.0909	0	.9062	.0428	.000271
			.0001	.9063	.0435	.000271 MIN
			.0005	.9065	.0462	.000271
			.0010	.9069	.0491	.000273
			.0050	.9081	.0660	.000295
. .	.04	.0909	0	.9059	.0407	.000347
			.0001	.9060	.0415	.000346 MIN
			.0005	.9063	.0444	.000347
			.0010	.9066	.0473	.000351
			.0050	.9079	.0650	.000384
			.0100	.9085	.0801	.00418
			.0150	.9086*	.0923	.000444
			.0200	.9084	.1029	.000465
. .	.08	.0909	0	.9045	.0312	.000565
			.0001	.9047	.0327	.000571
			.0005	.9053	.0375	.000583
			.0010	.9058	.0418	.000593
			.0050	.9073	.0610	.000699
.2	.05	.167	.0001	.8250	.0884	.000783
			.0005	.8253	.0898	.000781

TABLE 1—(Continued)

s	σ_s^2	x	u	$E(W)$	$\bar{x}$	σ_x^2
			.0010	.8256	.0916	.000778
			.0050	.8277	.1038	.000772 MIN
			.0100	.8294	.1163	.000774
			.0150	.8304	.1270	.000779
			.0200	.8313	.1367	.000784
			.0250	.8319	.1455	.000788
			.0300	.8322*	.1537	.000791
			.0600	.8306	.1927	.001568
	.10	.167	.0001	.8224	.0794	.001464
			.0005	.8228	.0813	.001457
			.0010	.8233	.0836	.001452 MIN
			.0050	.8259	.0977	.001455
			.0100	.8279	.1113	.001476
			.150	.8292	.1227	.001497
			.0200	.8301	.1328	.001515
			.0300	.8312	.1505	.001542
			.0400	.8315*	.1659	.001558
			.0450	.8314	.1731	.001566
.3	.10	.231	0	.7532	.1292	.001985
			.0001	.7533	.1295	.001983
			.0005	.7536	.1306	.001977
			.0010	.7540	.1320	.001969
			.0050	.7565	.1420	.001929
			.0100	.7589	.1529	.001898
			.0200	.7622	.1715	.001860
			.0400	.7658	.2018	.001805
			.0600	.7669	.2271	.001752
	.20	.231	0	.7463	.1129	.003760
			.0001	.7464	.1134	.003753
			.0050	.7469	.1151	.003729
			.0010	.7476	.1171	.003707
			.0050	.7513	.1301	.003636
.4	.05	.286	0	.6983	.1842	.001135
			.0001	.6983	.1844	.001134
			.0005	.6985	.1851	.001132
			.0010	.6988	.1860	.001129
			.0050	.7007	.1932	.001106
			.0100	.7028	.2015	.001083
			.0200	.7062	.2168	.001044
			.0400	.7104	.2433	.000984
			.0600	.7125	.2664	.000934
	.10	.286	0	.6951	.1787	.002272
			.0001	.6951	.1789	.002271
			.0005	.6954	.1797	.002265
			.0010	.6957	.1807	.002258
			.0050	.6979	.1884	.002209
	.20	.286	0	.6880	.1663	.004524
			.0001	.6881	.1665	.004520
			.0005	.6884	.1676	.004503

			.0010	.6889	.1688	.004484
			.0050	.6920	.1781	.004367
			.0100	.6950	.1884	.004265
			.0400	.7053	.2356	.003898
			.0600	.7082	.2606	.003715
			.0700	.7089	.2717	.003629
			.0750	.7091	.2771	.003586
			.0800	.7093	.2823	.003540
			.0850	.7094*	.2875	.003504
			.0900	.7093	.2926	.003463
	.30	.286	0	.6797	.1514	.006675
			.0001	.6798	.1518	.006664
			.0005	.6805	.1533	.006622
			.0010	.6812	.1551	.006578
			.0050	.6856	.1670	.006377
			.0100	.6896	.1789	.006238
			.0200	.6952	.1987	.006050
			.0400	.7019	.2303	.005770
			.0800	.7068	.2790	.005272
.8	.05	.444	0	.5471	.3840	.001050
			.0001	.5471	.3841	.001049
			.0005	.5472	.3846	.001047
			.0010	.5473	.3851	.001044
			.0050	.5483	.3894	.001024
	.10	.444	0	.5449	.3825	.002089
			.0001	.5449	.3826	.002088
			.0005	.5440	.3830	.002083
			.0010	.5451	.3835	.002077
			.0050	.5461	.3878	.002028
			.0100	.5473	.3930	.001966
			.0200	.5491	.4030	.001843
	.20	.444	0	.5401	.3766	.003975
			.0001	.5401	.3767	.003972
			.0005	.5403	.3771	.003959
			.0010	.5404	.3776	.003943
			.0050	.5416	.3818	.003819
			.0100	.5430	.3867	.003668
			.0200	.5453	.3961	.003376
			.0400	.5486	.4124	.002833
			.0600	.5507	.4261	.002347

It is apparent from the table that in the absence of mutation the average gene frequency is less than the optimum, that it rises with mutation to the lethal, and that the fitness is increased thereby. The rate of increase of $\bar{x}$ with mutation is quite large, as high as 15, while the slope of fitness against mutation is about 1 for the interval (0, .001) and then decreases. For some parameters we calculated optimum mutation rates which turn out to be approximately the values required to bring $\bar{x}$ up to $\hat{x}$. The optimum mutation increases with the variance of the environment. However, all optima were found to be very much greater than observed mutation rates. Possible reasons for this are considered in the DISCUSSION.

The variance of x has a more complex behavior. For $\bar{s} = .05$, it increases with

mutation. For $\bar{s} = .10$ or .20 and σ_s^2 small, the variance of x decreases to a minimum and then increases. For larger variances, or for $\bar{s}$ greater than .20, the variance of x decreases with mutation rate. But it is generally quite small and changes slowly with mutation, so that effectively the mean determines the optimum mutation rate.

Model II. Quadratic deviation model

If mutation is allowed to occur in both directions in such a way that equilibrium under mutation alone is the same as the mean gene frequency under selection alone, then mutation will not change the mean and we can study its effects on other components of fitness. For this purpose we consider the quadratic deviation model shown in Table 2. This model has been used in previous papers of this series (LEVINS 1964a,b, 1965) to study other aspects of adaptive processes.

The fitness of a population at any time is

2.01 $$\overline{W} = 1 - (s-M)^2 - \text{VAPHE}$$

where s is the environmentally determined optimum phenotype, M is the mean phenotype of the population, and VAPHE is the phenotypic variance within the population and is therefore $2a^2\, x\, (1-x)$. Since s is a random variable, there is no genetic equilibrium. But a steady state distribution is possible, and will give an average fitness

2.02 $$E(\overline{W}) = 1-\sigma_s^2 -(\bar{s}-\overline{M})^2 - \sigma_M^2 - E(\text{VAPHE}) + 2\,\text{COV}(s,M).$$

$\text{COV}(s,M)$ is the covariance of the environmental variable s and the mean phenotype of the population. Since the present phenotypic mean is the result of natural selection in the past, M is correlated with past values of s. Therefore if successive values of s are correlated among themselves, M is also correlated with the current s, and $\text{COV}(s,M)$ is positive.

This holds for a quadratic deviation fitness model in general. When the model of Table 2 holds, phenotype is determined by a single locus with two alleles, no dominance on the phenotype scale, and an additive phenotypic effect a. Then the components of fitness can be expressed in terms of gene frequencies, x. VAPHE is $2a^2x(1-x)$, so that its expected value is $2a^2\bar{x}(1-\bar{x}) - 2a^2\sigma_x^2$. The variance of the mean phenotype is $4a^2\sigma_t^2$. We have chosen to use a symmetric model in which $s = 0$, $M = 0$, $\bar{x} = .5$. Then the average fitness becomes

2.03 $$E(\overline{W}) = 1-\sigma_s^2 - a^2/2 - 2a^2\sigma_x^2 + 4a\,\text{COV}(s,x)$$

TABLE 2

The quadratic deviation model

	Genotype		
	AA	*Aa*	*aa*
Phenotype	a	0	$-a$
Frequency	x^2	$2x(1-x)$	$(1-x)^2$
Fitness	$1-(s-a)^2$	$1-s^2$	$1-(s+a)^2$

It was shown in the previous papers that if the correlation between the environments of successive generations is great enough (> 0.8) then the optimum value of a is different from zero, and the response to selection increases fitness. For any value of a, the rate of change of gene frequency can be approximated by the continuous process

$$2.04 \qquad \frac{dx}{dt} = a^2x(1-x)(1-2x) + 2asx(1-x) + u(1-2x)$$

where u is the mutation rate, taken to be equal in both directions. Since mutation always pushes the gene frequency toward its mean value of .5, the variance of x (and hence of the mean phenotype) is decreased. But VAPHE has its greatest value when $x = 0.5$. Hence, symmetric mutation increases VAPHE and decreases the variance of the mean. The latter effect is twice the former,

$$\text{since} \qquad \frac{\partial\, Var(M)}{\partial\, \sigma_x^2} = 4a^2 \qquad \text{while} \qquad \frac{\partial\, \text{VAPHE}}{\partial\, \sigma_x^2} = -2a^2$$

Thus the overall effect of mutation is to increase fitness.

Owing to the special assumption of symmetry, the only effects of mutation are on σ_x^2 and $\text{COV}(s,M)$. If the environments of different times are uncorrelated, mutation only reduces σ_x^2 and is thus unconditionally advantageous. Any asymmetry in the model (either in mutation rates that are unequal in opposite directions or in $\bar{s} \neq 0$) would result in mutation displacing $\bar{x}$ from its optimum. This could offset the effect on the variance and result in a low optimum mutation rate.

In the symmetric model with correlation among environments, it was shown in the third paper of this series that

$$2.05 \qquad \text{COV}(s,x) = 2a\alpha\sigma_s^2\,(1/4 - \sigma_x^2 + 6a\alpha\sigma_x^2).$$

Here α is defined from the correlation between two environments, a large α indicating strong correlation:

$$2.06 \qquad Cor(s_t, s_{t+h}) = e^{-h/\alpha}$$

Mutation increases $\text{COV}(s,x)$ by increasing the correlation of s with x. This happens by way of a damping action. Mutation is constantly pulling the gene frequency back toward its mean. Thus at any time the gene frequency depends more closely on the environments of the recent past while the environments of the remote past have a reduced contribution. However, mutation also reduces σ_x^2 so that for large values of u, $\text{COV}(s,x)$ will actually fall.

Hence for a symmetric model and no correlations among environments, an infinite mutation rate is advantageous. The introduction of correlations in the environment reduces the optimum mutation rate although it increases the contribution of that mutation rate to fitness. Finally, asymmetry in the model drastically reduces the optimum mutation rate.

These qualitative arguments can be made somewhat more precise by numerical methods. We used the same Monte Carlo simulation program on the IBM 7074 which was described in LEVINS (1965). In each generation a random s was generated with the appropriate statistical properties and used in the equation for genetic change

2.07 $$\Delta x = \frac{\{a^2x(1-x)(1-2x) + 2asx(1-x) + u(1-2x)\}}{\overline{W}}$$

which is more exact than the continuous expression in 2.04 for discrete generations. The components of fitness were calculated from ten replicate runs of 100 generations each for different environmental and genetic parameters. Since values of a greater than zero are advantageous for autocorrelations greater than about .8, we used values of .8, .9, and .99. Most work was done with an environmental variance of .1, which was chosen to give a stable distribution. The vaules of a, the additive phenotypic effect were taken near their optimum values.

In Table 3 we show the optimal mutation rates for various parameters. Mutation increases the average phenotypic variance in the steady state, and the magnitude of this increase relative to the original phenotypic variance is shown in the table to range up to about one third of the variance. The increase in VAPHE due to new mutation is of course less. This can be found as follows:

2.08 $$\frac{\partial \text{ VAPHE}}{\partial u} = 2a^2\,(1-2x)\,\frac{\partial x}{\partial u}$$

and for the change in one generation

2.09 $$\frac{\partial x}{\partial u} = 1-2x$$

so that

2.10 $$\frac{\partial \text{VAPHE}}{\partial u} = 2a^2\,(1-2x)^2$$

TABLE 3

Optimum mutation rates and variance due to mutation at optima. In 2-locus cases, recombination is 50% and ρ is the environmental autocorrelation

Number of loci	ρ	σ_s^2	a	$\hat{u}$	VAPHE ($\hat{u}$) — VAPHE (o) / VAPHE (o)	Proportion of New variance due to mutation
1	.8	.1	.12	.018	.22	2×10^{-3}
	.	.	.24	.035	.29	7×10^{-3}
	.9	.1	.06	$<10^{-4}$	~0	~0
	.	.	.12	.007	.19	1.6×10^{-3}
	.	.	.24	.016	.34	2×10^{-3}
	.9	.025	.08	.004	.02	....
	.	.050	.08	.006	.05	....
	.	.10	.08	.008	.12	....
	.	.20	.08	.016	.26	....
	.99	.1	.20	$<10^{-6}$	~0	~0
2	.8	.1	.05	.003	.03	3×10^{-4}
	.	.	.10	>.01	..	....
	.9	.1	.03	$<10^{-4}$	~0	~0
	.	.	.09	.002	.06	4×10^{-4}
	.	.	.08	.008	.12	1.7×10^{-3}
	.	.	.10	.010	.14	2.6×10^{-3}
	.	.	.12	.014	.19	3×10^{-3}

This has the expected value

2.11 $$E\{2a^2(1-2x)^2\} = 2\sigma_M^2$$

Thus the newly created variance is the fraction $2u\sigma_M^2$/VAPHE of the total VAPHE, and is shown in the last column of Table 3. We see that although the optimal mutation rate per locus is quite high, the new variance created by mutation is in all cases less than 1% of the total phenotypic variance.

We see from the table that the optimum mutation rate increases with a and with the environmental variance and decreases with the autocorrelation of the environment.

In Figure 1 we show fitness plotted against the mutation rate for various parameters, and in Figure 2 the effect of mutation on the covariance and correlation is shown. We see that although mutation increases fitness this effect is small, at best resulting in an increase of less than 1%. However, the contribution of the covariance to fitness can exceed this amount.

There is an element of arbitrariness in the model due to scaling, since we used the expression $W = 1 - (s - \text{phenotype})^2$. Fitness is lost owing to environmental fluctuation, and a portion of this fitness loss is restored by the response to selection. Since the fitness loss is on the order of 0.1, the environmental variance, the frac-

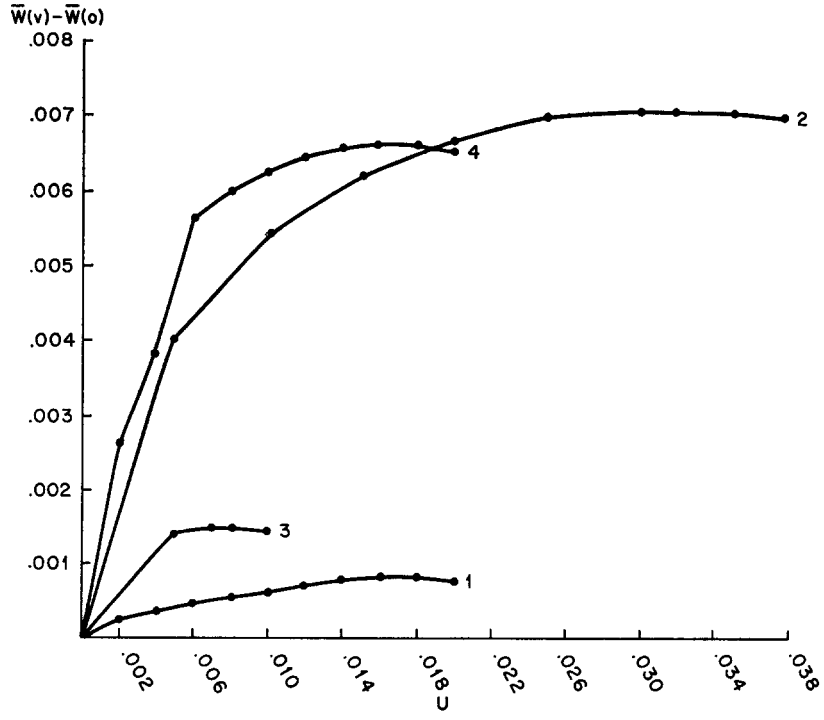

FIGURE 1.—Increase in average fitness as a function of mutation rate. The abscissa is mutation rate in both directions, the ordinate is $E(\bar{W})$ minus the average fitness for no mutation. Curve 1: $a = .12$, $p = .8$. Curve 2: $a = .24$, $p = .8$. Curve 3: $a = .12$, $p = .9$. Curve 4: $a = .24$, $p = .9$.

tion restored by mutation is on the order of 10%. With a large phenotypic variance, an autocorrelation in the environment of .9 can result in a correlation of gene frequency and environment of .5.

DISCUSSION

In both models, optimal mutations have been calculated which are very much greater than the per-locus mutation rates normally observed in nature. This discrepancy suggests three possibilities: (1) Our calculations in model II are based on a single locus. But most quantitative characters are polygenic. The optimum mutation rate must be expressed in terms of variance per unit phenotype instead of per locus. (2) Genes which affect mutation rate would have quite small selective advantage unless they affected mutation at many loci. But not all loci satisfy the requirements of our models. At most loci, mutation may

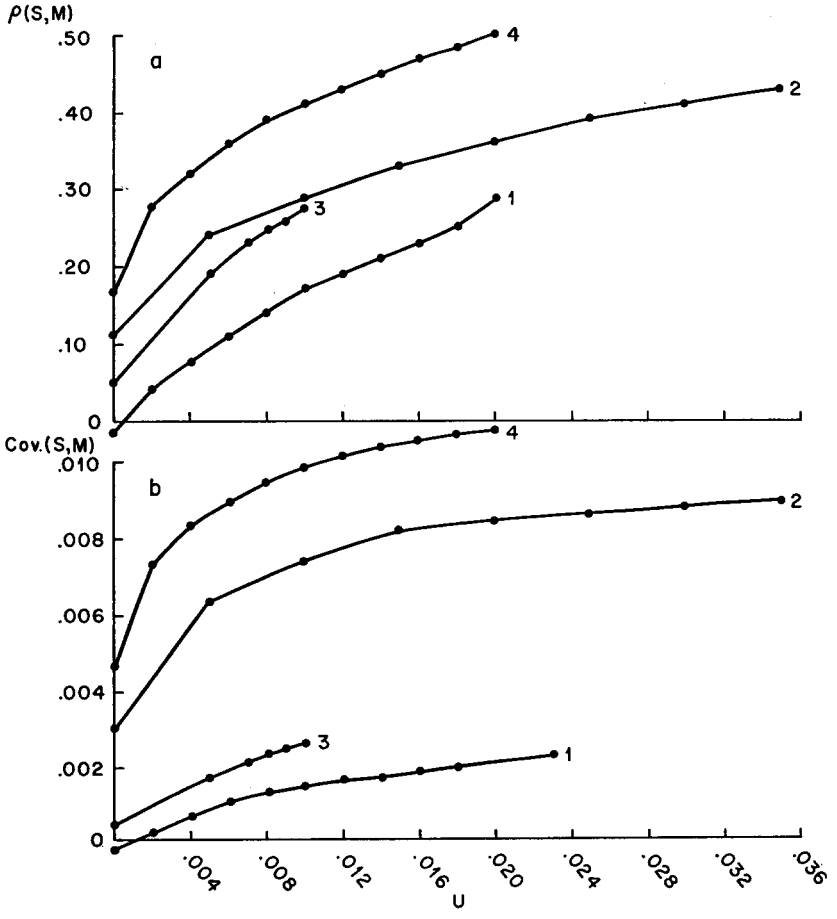

FIGURE 2.—(a) Correlation between the mean phenotype and the optimum phenotype s. (b) The covariance of mean phenotype and the optimum phenotype s. Curve 1: $a = .12$, $p = .8$. Curve 2: $a = .24$, $p = .8$. Curve 3: $a = .12$, $p = .9$. Curve 4: $a = .24$, $p = .9$.

be harmful, so that the optimum mutation rate will be a compromise between the optima at relatively few heterotic loci with variable adaptive values of the viable homozygote, and the optimum of zero at the rest of the loci. (3) That a given mutation rate is advantageous does not guarantee that it will be selected. We are dealing here with a second order type of selection in which the gene in question does not appear directly in the expression for $\overline{W}$ but acts only through its effect on the frequency of other genes.

Polygenic quantitative characters: In the quadratic deviation model we considered a phenotype controlled by two additive loci contributing with effects a_1 and a_2, respectively. For $p = .9$, $\sigma_s^2 = .1$, the optimum $a_1 + a_2 = .12$. The loci were considered to segregate independently. In Table 4 we show the optimum mutation rates for different partitioning of the phenotype between the loci. It is clear that the mutation rate per locus, the total mutation rate over two loci, and the increase in phenotypic variance due to mutation are all reduced when the additive phenotypic effect is spread more evently. It was also found that if the phenotype was controlled by two loci but mutation restricted to one locus, the optimum mutation rate at that locus was greater. A locus with a relatively small part of the total phenotype required a much higher mutation rate if it had to provide all the mutational variance.

Our computing system did not permit any extension to large numbers of loci. However, it is already clear that polygenic systems in our model have lower optimum mutation rates than single locus quantitative genes.

Combined optima: For most loci, the effect of mutation is to reduce fitness by an amount equal to the mutation rate. This is the familiar mutation load. For the loci of models I and II, fitness increases with the mutation rate with a slope that decreases as u increases. Call this function $W(u)$. If the proportion p of the loci are of this type, and $1-p$ exhibit ordinary load, then the total fitness for genes with the same mutation rate will be $pW(u) - (1-p)u$. It will continue to increase with u as long as

$$3.01 \qquad \frac{dW(u)}{du} > \frac{(1-p)}{p}$$

TABLE 4

Optimum mutation rates and increased variance at optimum for two loci

a_1	a_2	$\hat{u}$	VAPHE ($\hat{u}$) — VAPHE (o) / VAPHE (o)
.12	0	.007	.19
.11	.01	.006	.20
.09	.03	.004	.11
.07	.05	.002	.05
.06	.06	.002	.04

a_1 and a_2 are the phenotypic effects of the two loci. The last column is the proportionate increase in variance at optimum mutation rate over that at no mutation. The autocorrelation of the environment is .9 and the variance .1. Recombination is 50%.

From Table 1 we see that in all cases dW/du falls below 2 for some $u < .0001$. Therefore, unless more than a third of the loci are of the type used in model I, the combined optimal mutation will be below 10^{-4}. And if dW/du is always less than $(1-p)/p$, the optimum will be zero. The slope of $W(u)$ can be found from 1.05.

Recalling our assumption that there are no correlations among successive environments (so that $E(sx) = \bar{s}\bar{x}$) and that $\hat{x} = \bar{s}/(1+\bar{s})$, we have

3.02 $$E(\overline{W}) = 1-\hat{x} - (1+\bar{s})(\hat{x}-\bar{x})^2 - (1+\bar{s})\,\sigma_x^2$$

Differentiating with respect to the mutation rate u gives

3.03 $$\frac{\partial E(W)}{\partial u} = 2(1+\bar{s}(\hat{x}-\bar{x})\frac{\partial \bar{x}}{\partial u} - (1+\bar{s})\frac{\sigma \partial_x{}^2}{\partial u}$$

We can find $\partial\bar{x}/\partial u$ from the equation for the rate of change of the mean gene frequency, which is the same as 1.07 with the term $(1-x)u$ added to show mutation toward the lethal:

3.04
$$\frac{d\bar{x}}{dt} = \tfrac{1}{2}\{x(1-\bar{x})(\hat{x}-\bar{x})(1+\bar{s}) + [3\hat{x}(1+\bar{s}) - (2\bar{s}+1)]\sigma_x^2 + (1+\bar{s}(u_3\} + (1-x)u$$

At the steady state, $d\bar{x}/dt = 0$. The right side can then be differentiated with respect to the mutation rate u. The higher moments change only very slowly with u, as seen from Table 1. Thus we find that near $u = 0$

3.05 $$\frac{\partial \bar{x}}{\partial u} = \frac{2(1-\bar{x})}{\{(1-2\bar{x})(\hat{x}-\bar{x}) - \bar{x}(1-\bar{x})\}(1+\bar{s}) + 3\sigma_x^2(1+\bar{s})}$$

and finally

3.06 $$\frac{\partial E(\overline{W})}{\partial u} = \frac{4(\hat{x}-\bar{x})(1-\bar{x})}{(1-2\bar{x})(\hat{x}-\bar{x}) - \bar{x}(1-\bar{x}) + 3\sigma_x^2}$$

The values of $\hat{x}$, $\bar{x}$, and σ_x^2 can be found from Table 1. The most favorable case occurs for $\bar{s} = .1$, $\sigma_s^2 = .04$. There the estimated slope $\partial E(\overline{W})/\partial u$ is about 24.

Thus, if even only 4% of the loci are of this type, there will be an optimum mutation rate greater than zero. But $\partial E(\overline{W})/\partial u$ decreases rapidly with u, and is less than 1 at $u = 10^{-4}$. Thus the optimum will certainly be less than 10^{-4}.

Selection for the mutation rate: A gene whose only effect is to alter the mutation rate does not appear directly in the expression for $\overline{W}$. However, it will not be in linkage equilibrium with the gene whose mutation it affects. Therefore selection on the principal locus will carry selection for mutation rate along with it. Consider an ordinary locus with the alleles X_1, X_2 and the genotypic fitnesses W_{11}, W_{12}, W_{22}. In the absence of mutation, selection will carry such a locus to equilibrium at a frequency of X_1 given by

4.01 $$\hat{x} = \frac{(W_{12}-W_{22})}{\{2W_{12}-W_{11}-W_{22}\}}$$

We do not require that this be a stable polymorphism. Mutation may occur in both directions at this locus. Under mutation alone (in the absence of selection) x would reach an equilibrium value at

4.02 $$x_m = u/(u+v)$$

where u and v are the mutation rates to and from x_1. Since the mutation rate is influenced by a second locus, x_m may be a function of the frequency y of the mutation rate gene Y_1.

When x is less than x_m, increased mutation increases x. Thus, for those gametes which carry Y_1 the frequency of X_1 is $x+e$, where e is positive for $x < x_m$ and negative for $x > x_m$. The value of e also depends on the closeness of linkage between the X and Y loci. Similarly, the gametes carrying Y_2 have a reduced frequency of X_2, equal to x-f. Since the frequency of X_1 is x, $ye = (1-y)f$. In Table 5 we show the gamete frequencies, genotype frequencies, and their fitnesses.

From the table we can calculate the marginal fitness of the Y_1 allele, which is equal to the weighted average of its homozygous and heterozygous fitnesses. This is

4.03 $$W(Y_1) = \overline{W} + e\,[W_{12}-W_{22}-(2W_{12}-W_{11}-W_{22})x]$$

Thus it follows that y will increase whenever x_m-x and $\hat{x}-x$ have the same sign.

The graphical representation in Figure 3 permits us to follow the joint changes in x and y. Under the influence both of selection and mutation, x will approach some value between $\hat{x}$ and x_m. The higher the mutation rate, the closer it will be to x_m. This equilibrium value is shown in the graph by the dotted line. Meanwhile, we see from 4.03 that y increases whenever x is outside the interval $(\hat{x},x_m)$ and decreases when x lies within that interval. The results are shown by the arrows in Figure 3. We see that in a constant environment, selection may initially favor an increased mutation rate but eventually x will enter the interval $(\hat{x},x_m)$ and selection will reduce mutation. If $\hat{x} = x_m$, then of course, mutation rate will increase until $y = 1$, but this is infinitely improbable except when both are equal to 0 or 1. In Figure 3b we show that if mutation rate is not

TABLE 5

Model for selection of a gene increasing mutation rate

Gamete	Frequency
Y_1X_1	$y(x+e)$
Y_1X_2	$y(1-x-e)$
Y_2X_1	$(1-y)(x-f)$
Y_2X_2	$(1-y)(1-x+f)$
Genotype	**Fitness**
Y_1Y_1	$W_{11}(x+e)^2 + 2W_{12}(+e)(1-x-e) + W_{22}(1-x-e)^2$
Y_1Y_2	$W_{11}(x+e)(x-f) + 2W_{12}[(x+e)(1-x+f) + (1-x-e)(x-f)] + W_{22}(1-x-e)(1-x+f)$
Y_{22}	$W_{11}(x-f)^2 + 2W_{12}(x-f)(1-x+f) + W_{22}(1-x+f)^2$

Marginal fitness of $Y_1 = yW(Y_1Y_1) + (1-y)W(Y_1Y_2)$

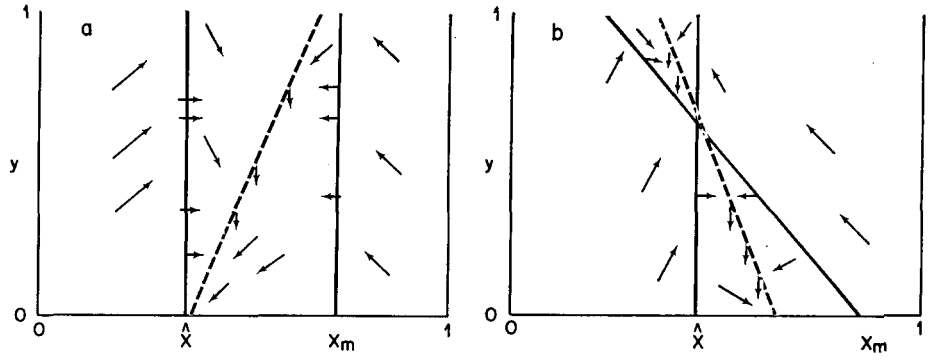

FIGURE 3.—Selection for mutation rate. The abscissa is the frequency x of the principal gene, the ordinate is the frequency y of the gene that increases mutation. The dotted line is the equilibrium value of x for each y, x is the equilibrium of $\hat{x}$ under selection alone and x_m is the equilibrium of x under mutation alone. Selection and mutation together move x toward the dotted line while y increases outside the interval $(\hat{x},x_m)$ and decreases within the interval.

affected equally in both directions, for some value of y we may have $x = x_m$. This does not change our conclusion.

In a variable environment, $\hat{x}$ is no longer constant, and x is distributed around some mean value. In model I, mutation was toward increased x, and we showed that x is usually below the average $\hat{x}$. Thus selection will favor an increased mutation rate. But as the mutation rate approaches its optimum, the average gene frequency approaches $\hat{x}$. Then x is above and below $\hat{x}$ more or less equally often, and progress halts.

In model II, the symmetry assumptions are such that without mutation, or with $x_m = \hat{x}$, x is equally often above or below $\hat{x}$. If x_m is much greater than $\hat{x}$, x is usually between the two and mutation rate will be reduced. But if x_m is near enough to .5, and if the variance of the environment is great enough, then x will lie outside the interval most of the time and an increased mutation rate may be favored, or a polymorphism for mutation rate genes may be maintained.

The models used in this study are of course rather specialized. They were chosen not because they represent common situations but because they emphasize particular aspects of the effects of mutation. The essential feature of the first model is that under fluctuating conditions the frequency of the least favored allele will be below optimum. If mutation toward a lethal is advantageous, it would also be advantageous in models with less deleterious homozygotes.

Similarly, model II isolates the effects of mutation on variance of gene frequency and on the correlation between gene frequency and environment.

Although both models show ways in which recurrent mutation contributes to fitness, the optimal values which they predict are excessive. But we cannot correct these estimates quantitatively on the basis of present knowledge. We do not know what proportion of the loci influenced by a given mutation rate gene are of the type required (with unequal average fitnesses of homozygotes and fitnesses fluc-

tuating widely with the environment). We don't know the variance of the selection coefficients found in nature. We do not know over how many loci the phenotypic effects are spread. Therefore, the theory can only lead to qualitative predictions in the form of inequalities:

(1) Both models predict that mutation rates will be greater in variable than in constant environments. This could be tested comparing temperate and tropical populations of the same or related species. (2) Widespread, ecologically plastic species experience greater environmental diversity than those with narrow specialized niches. Thus we expect the mutation rate to be greater in the species with broad adaptation. (3) Within species, the mutation rates should be greater for phenotypes whose relative fitness is very sensitive to environmental change, and less for traits whose fitness is more fixed, such as those related to the canalization system or reproductive organs. Mutation rate here must be measured as rate of production of phenotypic variance. (4) For loci of variable fitness, mutation toward the allele of lower average fitness should be greater. (5) It may be possible to demonstrate selection for mutation rate in laboratory populations by the appropriate pattern of varying selection on the phenotype.

SUMMARY

The effects of mutation rate on the components of population fitness in a variable environment were studied anlytically for populations with the steady state distribution of gene frequency and by semi-Monte Carlo simulation on a computer. Population fitness was analyzed into components due to changes in the mean gene frequency, the variance, and the correlation between gene frequency and environment. In an asymmetric model using a single heterotic locus it was shown that the mean gene frequency of the less favored allele under variable selection is below optimum, so that mutation toward this allele is advantageous. In a symmetric quadratic deviation model, the advantage of mutation is the reduction of variance of gene frequency and an increase in the correlation between the mean phenotype of the population and the optimum phenotype which is an environmental variable. However mutation had the adverse effect of increasing the average phenotypic variance within the population at any given time. These different effects of mutation result in optimal mutation rates which were calculated to be much higher than those observed in nature. However, when the effects are spread over many loci and we look at the total phenotypic variance added by the optimal rates it may fall within the observed range. It was shown that selection may increase the mutation rate in a variable environment especially in asymmetric models. Finally, it is concluded that mutation rates in nature should be greater in broad-niched, unspecialized species than in restricted species, in variable climates than in stable ones, and for traits whose fitness is sensitive to the environment than in traits of more or less constant fitness.

LITERATURE CITED

HALDANE, J. B. S., 1957 The cost of natural selection. J. Genet. **55**: 511–524.

KIMURA, M., 1960 Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load. J. Genet. **57**: 21–34.

LEVINS, R., 1964a Theory of fitness in a heterogeneous environment. III. The response to selection. J. Theor. Biol. **7**: 224–240. —— 1964b Theory of fitness in a hetergoeneous environment. IV. The adaptive significance of gene flow. Evolution **18**: 635–638. —— 1965 Theory of fitness in a heterogeneous environment. V. Optimum genetic systems. Genetics **52**: 891–904.

WRIGHT, S., 1931 Evolution in Mendelian populations. Genetics **16**: 97–159.