THE MUTATION LOAD IN SMALL POPULATIONS[1]

MOTOO KIMURA[2], TAKEO MARUYAMA, and JAMES F. CROW

University of Wisconsin, Madison, Wisconsin

Received April 29, 1963

THE mutation load has been defined as the proportion by which the population fitness, or any other attribute of interest, is altered by recurrent mutation (MORTON, CROW, and MULLER 1956; CROW 1958). HALDANE (1937) and MULLER (1950) had earlier shown that this load is largely independent of the harmfulness of the mutant. As long as the selective disadvantage of the mutant is of a larger order of magnitude than the mutation rate and the heterozygote fitness is not out of the range of that of the homozygotes, the load (measured in terms of fitness) is equal to the mutation rate for a recessive mutant and approximately twice the mutation rate for a dominant mutant. A detailed calculation of the value for various degrees of dominance has been given by KIMURA (1961).

In all these studies it has been assumed that the population is so large and the conditions so stable that the frequency of a mutant gene is exactly determined by the mutation rates, dominance, and selection coefficients, with no random fluctuation. However, actual populations are finite and also there are departures from equilibrium conditions because of variations in the various determining factors. Our purpose is to investigate the effect of random drift caused by a finite population number.

It would be expected that the load would increase in a small population because the gene frequencies would drift away from the equilibrium values. This was confirmed by our mathematical investigations, but two somewhat unexpected results emerged. One is that, for a given population size, a mildly deleterious mutant may create a considerably larger load than a more deleterious one. The second is that, under some circumstances, a finite population may have a smaller load than an infinite one.

MATHEMATICAL METHODS

We consider the following model, where A' may be regarded as a single mutant allele or a class of mutant alleles with collective frequency x. Random mating is assumed within a finite population of effective size N. The average fitness, W, of each genotype is expressed as a fraction of the fitness of AA.

[1] Paper No. 918 from the Division of Genetics; also Contribution No. 473 from the National Institute of Genetics, Mishima-shi, Japan. Sponsored by the Mathematics Research Center, U. S. Army, University of Wisconsin, (Contract No. DA-11-022-ORD-2059) and supported in part by the National Institutes of Health (RG-7666).

[2] On leave from the National Institute of Genetics, Mishima-shi, Japan.

Genotype:	AA	AA'	$A'A'$
Frequency:	$(1-x)^2$	$2x(1-x)$	x^2
Average fitness, W:	1	$1-hs$	$1-s$

Mutation rate (forward), $A \longrightarrow A' = u$
Mutation rate (reverse), $A' \longrightarrow A = v$

$$\overline{W} = 1 - 2hsx(1-x) - sx^2 \tag{1}$$

Various degrees of dominance are included by varying h; when $h = 0$, the mutant gene is recessive, and when $h = \frac{1}{2}$, there is no dominance. The probability distribution of x is given by WRIGHT's (1937) basic equation

$$\phi(x) = C\overline{W}^{2N} x^{4Nu-1}(1-x)^{4Nv-1} \tag{2}$$

and the mutation load of a population with mutant gene frequency x is

$$L = 1 - \overline{W} = 2shx(1-x) + sx^2 \tag{3}$$

where $s > 0$ and $h \geq 0$. (To avoid possible misunderstanding (see LI 1963) it should be emphasized that the fitnesses assigned to the various genotypes are relative, not absolute. More precisely, $L = (W_{max} - \overline{W})/W_{max}$, where W_{max} is the fitness of the most fit genotype, in this case AA. Since the fitnesses are relative, W_{max} is assigned the value 1 for algebraic convenience.)

The mean value of the contribution of this locus to the mutation load is then given by

$$\overline{L} = \frac{\int_0^1 L\phi(x)\,dx}{\int_0^1 \phi(x)\,dx} \tag{4}$$

The numerical evaluation of this integral causes some difficulties. For the case, $h = \frac{1}{2}$, the integrals can be expressed in terms of the confluent hypergeometric function if the fitness is measured in Malthusian parameters (see APPENDIX). In this and other cases, numerical integration by quadrature was mainly used. In some cases the integrals were replaced by summations based on a discrete model. These procedures are described in an appendix for the case of genic selection. Several cross checks were made by computing the same value in alternate ways. In some special cases the integrals can be expressed by elementary functions, and these also serve as checks on the numerical integration of more complex cases: (1) The case of a lethal gene without dominance ($h = \frac{1}{2}$, $s = 1$) can be expressed directly in terms of Beta and Gamma functions; (2) The case of a recessive lethal can also be expressed with Gamma functions by the transformation, $y = x^2$; (3) The special cases $4Nu = 4Nv = 1$, $h = \frac{1}{2}$, and $4Nu = 2$, $4Nv = 1$, $h = 0$ can be integrated directly.

For small values of Ns ($2Ns < 5$) the following approximation is useful:

$$\bar{L} = \frac{s}{1 + \frac{v}{u} e^{2Ns}} \tag{5}$$

At the other extreme, when Ns becomes very large, the mutation load is given approximately by u when $h = 0$ and by $2u$ when $h = \frac{1}{2}$. For other values of h,

$$\bar{L} = u(1 - \theta + \sqrt{\theta(2 + \theta)}), \tag{6}$$

$$\theta = sh^2/2u(1 - 2h), \qquad 0 < h \leq 0.5.$$

RESULTS

Numerical results with no dominance and complete dominance: Some approximate numerical values are given in graphical form in Figure 1. We take as a typical value of the forward mutation rate, $u = 10^{-5}$. Reverse mutation is usually a much slower process, so we have taken $v = 10^{-6}$. Figure 1 gives the average mutation load, $\bar{L}$ as a function of the effective population number, N. The lines

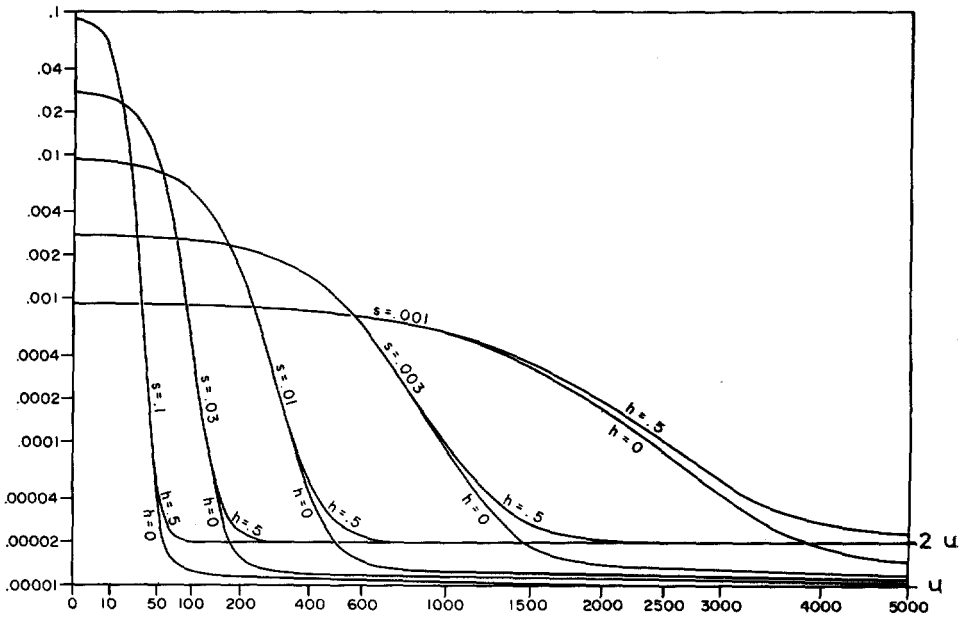

FIGURE 1.—The mutation load (ordinate) as a function of the effective population number (abscissa) for various values of the selection coefficient, s, and dominance, h. The ordinate is on a logarithmic scale; the abscissa is on a scale proportional to the square root. The forward mutation rate, u, is 10^{-5}; the reverse rate v, is 10^{-6}.

correspond to selection coefficients, $s = 0.1, .03, .01, .003$, and .001, and two levels of dominance, $h = .5$ (no dominance) and $h = 0$ (recessive mutant).

When the population is very small the mutation load is determined mainly by the selection coefficient s and by the ratio of the forward and backward mutation rates, u/v, as can be seen from equation (5). The degree of dominance is irrelevant. This is reasonable, for in very small populations the mutant gene tends to be either fixed or absent, so there are few heterozygotes.

In a very large population the HALDANE (1937) principle applies and the load is equal to the mutation rate for a recessive mutant and equal to approximately twice the mutation rate for a dominant mutant as long as s is considerably larger than u. A mutant that is partially dominant is more like a dominant than a recessive in this respect. For example, equation 6 shows that even with only 5 percent dominance ($s = .01$, $h = .05$) the load is 1.78 u, much closer to $2u$ than to u. The rate of reverse mutation, v, is unimportant.

Another point of interest that appears in Figure 1 is the very slow approach to the large population value for a recessive mutant. As first shown by WRIGHT (1937) and recently emphasized by ROBERTSON (1962), even a rather large population may be far from equilibrium for the frequency of a recessive mutant. This is reflected in Figure 1. For example, when $s = .01$ and $N = 3500$ the average mutation load is 1.11 u, 11 percent in excess of the value in an equilibrium population. On the other hand when $h = .5$, the corresponding departure from large population equilibrium value is reached in a population of about 500.

The most important feature that is revealed by Figure 1 is that in many populations a mutant with a small selective disadvantage causes a greater mutational load than a mutant with a greater harmful effect. As can be seen from Figure 1, in a population larger than about 250 a mutant with $s = .001$ has a larger load than one with $s = .01$; for example, when $N = 900$ and $h = 0$, the load is approximately 50 times as large for $s = .001$ as for $s = .01$.

This means that, if there are loci where s is in the general region of 10^{-3} or 10^{-4}, even a population with an effective number of several thousand may have gene frequencies far from the equilibrium values for an infinite population. This could create an appreciable genetic load. For example, in a population of effective size 1000 if there were 1000 loci producing mutants with $s = .001$, the mutation load would be roughly one half. This is equivalent to 25,000 loci with $u = 10^{-5}$ in an infinite population.

Partial dominance: It is likely that a great many mutants are nearly, but not completely, recessive. For example, the experimental value of h for newly induced lethal mutations in Drosophila is about .05. Figure 2 shows the curves for $h = 0, .05$, and .5 when $s = .01$. The figure shows what was emphasized earlier—that in a large population a mutant that is nearly recessive produces a load more similar to that of a dominant than of a fully recessive mutant.

The most interesting part of the curve is in the region around $N = 1000$. Here there is the paradox that a finite population has a smaller load than an infinite population, which would seem to imply that a random process produces a higher average fitness than a deterministic one. The explanation lies in the different load

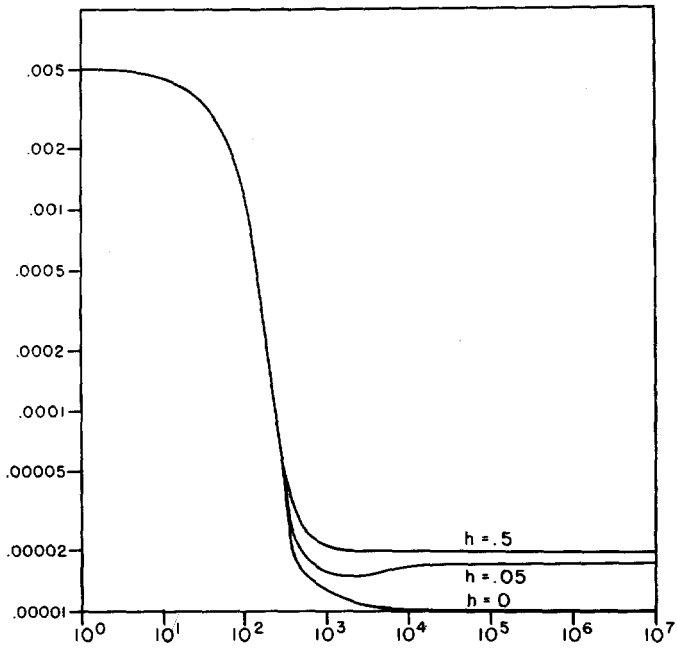

FIGURE 2.—The mutation load (ordinate) as a function of the effective population number (abscissa) for various levels of dominance, $h = 0$ (recessive mutant), $h = .05$, and $h = .5$ (no dominance). The selection coefficient, s, is .01 and the forward and reverse mutation rates are each equal to 10^{-5}.

produced by dominant and recessive factors. Heterozygous selection accounts for one gene elimination per genetic death; selection against homozygotes produces two gene eliminations per genetic death. In this case, with a mutant that is nearly recessive, the curve is more like that of a recessive for $N < 1000$, but then changes and becomes more like a dominant. This is because in a small population the gene is usually absent or fixed, in which case most of the mutant expression is in homozygotes and eliminations are as for a recessive mutant. With larger populations a large fraction have the mutant in intermediate frequencies so that it occurs more often in heterozygotes, and thus the small heterozygous effects become more important than the larger effects in the much rarer homozygotes.

In general, the mutation load is never less than the mutation rate (except for loci where the gene frequencies are maintained by balanced selective forces). In a large population it will usually be between one and two times the mutation rate, depending on the dominance. In some cases, probably rare, such as when the heterozygote is more deleterious than either homozygote, the load may be more than twice the mutation rate (KIMURA 1961). However, in a finite population the load may be many times the mutation rate, approaching the value of the selection coefficient s as a limit in small populations where the rate of reverse mutation is negligible.

The selection coefficient that maximizes the mutation load: We have shown in

the previous section that when two mutants having different selection coefficients are compared, there is a wide range of values of the effective population number where the gene with the smaller selective disadvantage produces the larger average load. In other words, individually less deleterious mutations cause more harm to the population in the long run. This result is understood if we note that the smaller the effect of a gene, the more its frequency is determined by random drift. The milder gene may be carried to a high frequency by random drift and thereby cause a larger effect than a more deleterious gene whose frequency is kept very low by selection.

For a given population size, there must therefore be a value of s which maximizes the average mutation load, $\bar{L}$. This may be obtained from equation (5). Letting $X = 2Ns$, we obtain from (5)

$$2N\bar{L} = \frac{X}{1 + \frac{v}{u}e^X} \tag{7}$$

If $u = 10v$, as in our previous examples, the value of X which maximizes $2N\bar{L}$ (and therefore $\bar{L}$) is approximately $X = 2.157$, for which $2N\bar{L}_{max} = 1.157$. For a population of size $N = 200$, the most damaging mutation is one with $s = .0054$, for which $\bar{L}_{max} = .0029$. For $N = 2000$, the mutant creating the largest mutational load would have $s = .00054$.

Multiple loci: Thus it would seem that slightly deleterious mutations of a type which may be very important for evolution may be very damaging to a small population when all the relevant loci are taken into account. If $\bar{L}_i$ is the mutation load at the ith locus, then the total fitness relative to a population with mutation rate zero would be

$$W_T = \Pi\ (1 - L_i) \backsim e^{-\Sigma L_i} \tag{8}$$

where the summation is over all relevant loci. The loci are assumed to be independent in their effect.

It is convenient to measure the mutation load in Malthusian parameters (FISHER 1930, 1958). In this case the total mutation load is

$$L_T = -\log W_T = \Sigma L_i \tag{9}$$

For example, for $N = 200$ and $s = .0054$, if there are 1000 such loci $\bar{L}_T = 1000 \times .0029 = 2.9$ and the relative fitness is $e^{-2.9} = 0.055$, a reduction of almost 95 percent. With 100 loci $L_T = .29$ and $e^{-.29} = .75$.

Interpopulation selection: The advantage of measuring the load by $-\log W_T$ rather than by $1 - W_T$ is that the total load becomes the sum of its components, rather than a complicated product expression. Since the L_i's are random variables, L_T will be distributed normally by the central limit theorem provided the number of loci is large. The mean and variance of L_T will be

$$\bar{L}_T = \Sigma\bar{L}_i,\ V_T = \Sigma V_i\ . \tag{10}$$

In a large group of populations having the same size and genetic parameters u, v, s, and h, L_T is distributed approximately normally with mean and variance given by (10). Since the relative selective value of a population having a partic-

ular value of total mutation load y is e^{-y}, the expected value of this quantity over all the populations is

$$\overline{W}_T = \overline{e^{-L_T}} = \frac{1}{\sqrt{2\pi V_T}} \int_{-\infty}^{\infty} e^{-y} e^{\frac{-(y-\bar{L}_T)^2}{2V_T}} dy = e^{-\bar{L}_T} e^{V_T/2} \tag{11}$$

Thus, the average fitness of a series of populations is a little greater (by a factor of $e^{V_T/2}$) than would be obtained by averaging the loads and computing the fitness from these.

The variance of the mutation load has an important bearing on the problem of intergroup selection. Table 1 gives the mutation load and its variance for a few values of s, h, and N.

As a numerical illustration, consider a group of 40 populations, each with 1000 individuals. Assume that there are 100 independent gene loci, each giving rise to mutations at rate $u = 10^{-5}$, with reverse mutation $v = 10^{-6}$, with $s = .001$, and no dominance ($h = 1/2$). From Table 1, $\bar{L} = 5.9 \times 10^{-4}$ and $V = 2.4 \times 10^{-7}$ for each locus; and for 100 loci $\bar{L}_T = .059$, $V_T = 2.4 \times 10^{-5}$, and the standard deviation of L_T is .0049. The best of the 40 populations would be expected to deviate about two standard deviations from the mean, hence would have an expected load of $.059 - 2(.0049) = .049$. This is to be compared to a load of $100 \times 2u = .002$ for a single population of 40,000. Even if the most fit population replaced all the others it would still come nowhere near to raising the fitness to that of a single panmictic population.

This means that if there are a number of independent loci subject to random drift, interpopulation selection is not very effective in raising the average fitness. Migration between the populations would have very little effect unless it were

TABLE 1

Mean (L) *and variance* (V_L) *of the mutation load in a population of effective number* N

s	h	N	L	V_L
.001	.5	100	8.9×10^{-8}	9.7×10^{-8}
		1000	5.9×10^{-4}	2.4×10^{-7}
		∞	2.0×10^{-5}	
.001	0	100	8.9×10^{-4}	9.7×10^{-8}
		1000	5.8×10^{-4}	2.4×10^{-7}
		∞	1.0×10^{-5}	
.01	.5	40	8.2×10^{-3}	1.5×10^{-5}
		400	5.6×10^{-5}	3.7×10^{-7}
		∞	2.0×10^{-5}	
.01	0	40	8.2×10^{-3}	1.5×10^{-5}
		400	4.8×10^{-5}	3.5×10^{-6}
		∞	1.0×10^{-5}	

The forward mutation rate, u, is 10^{-5} and the reverse mutation rate one tenth as large. Dominance is measured by h, and s is the selective disadvantage of the mutant homozygote.

at a high enough rate to convert the population into almost the equivalent of a single randomly mating unit. As WRIGHT has emphasized, the absolute amount of migration required for this may be quite small.

WRIGHT (1931 and later) has emphasized the idea that a subdivided population with local random differentiation, a small amount of migration, and intergroup selection has a favorable structure for breaking evolutionary stalemates and for creative evolution. Our analysis does not negate this possibility, but it does show that such a population structure pays a substantial price in reduced fitness. Any evolutionary advantage of such a system would have to overcome the initial disadvantage of having a considerably lower fitness than a large panmictic population.

SUMMARY

In a small population, the gene frequencies are subject to random fluctuations and the mutation load becomes a random variable. The distribution of gene frequencies under the influence of mutation and selection are known so the mean and variance of the mutation load can be computed.

In small populations, the load is considerably larger than in a large population. For a wide range of population sizes, a mutant that is slightly harmful is more damaging to the fitness of the population than a mutant with a much greater harmful effect. Intergroup selection is ineffective in reducing this load.

APPENDIX: NUMERICAL METHODS USED TO EVALUATE THE MEAN AND THE VARIANCE OF THE MUTATIONAL LOAD WITH GENIC SELECTION

MOTOO KIMURA

1. *Expression of the integrals in terms of the confluent hypergeometric function:* If we denote by $s/2$ the selective disadvantage of A' over its wild-type allele A measured in Malthusian parameters, then the probability distribution of the frequency x of A' is given by

$$\text{(A.1)} \quad \phi(x) = Ce^{-2Nsx}x^{4Nu-1}(1-x)^{4Nv-1}$$

where u is the forward mutation rate ($A \rightarrow A'$) and v is the reverse mutation rate ($A' \rightarrow A$). Note that if s is small $(1-sx)^{2N} = e^{-2Nsx}$ with good approximation and (A.1) is equivalent to equation (2) of the main text. We will denote by $P(x)$ the right side of (A.1) with C omitted, i.e.

$$\text{(A.2)} \quad P(x) = e^{-2Nsx}x^{4Nu-1}(1-x)^{4Nv-1}.$$

The average mutational load with respect to this locus is

$$\text{(A.3)} \quad \bar{L} = s\,{}_0\!\int^1 xP(x)dx \,/\, {}_0\!\int^1 P(x)dx.$$

This can be expressed as follows:

$$\text{(A.4)} \quad \bar{L} = \frac{su}{u+v} \cdot \frac{{}_1F_1(4Nu+1,\ 4Nu+4Nv+1,\ -2Ns)}{{}_1F_1(4Nu,\ 4Nu+4Nv,\ -2Ns)},$$

where ${}_1F_1$ denotes the confluent hypergeometric function defined by

$$(A.5) \quad {}_1F_1(\alpha, \gamma; z) = 1 + \frac{\alpha}{\gamma} \cdot \frac{z}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)} \cdot \frac{z^2}{2!} + \ldots$$

In the present case, $z = -2Ns$ and the above series was used to calculate $\overline{L}$ for Ns up to 10, above which an asymptotic formula for $|z| \to \infty$ was used to calculate the ${}_1F_1$'s and therefore $\overline{L}$.

Similar expressions may be obtained for $\overline{L^2}$ and the variance of the mutational load was calculated from

$$(A.6) \quad V_L = \overline{L^2} - \overline{L}^2.$$

II. *Numerical integration by quadrature:* Since for a small population number, the value of $P(x)$ may become $+\infty$ both at $x = 0$ and 1, caution is necessary in applying the method of quadrature in order to compute the values of integrals in (A.3). However, if we define function $Q(x)$ by

$$Q(x) = P(x) - x^{\alpha-1} - e^z(1-x)^{\beta-1},$$

where $\alpha = 4Nu$, $\beta = 4Nv$, $z = -2Ns$ and $P(x)$ being given by (A.2), then $Q(x)$ is always finite over the entire range of integration. Here we assign values $-e^z$ and -1 respectively to $Q(0)$ and $Q(1)$ in computation. Therefore the ordinary method of quadrature is applicable to compute the value of

$${}_0\int^1 Q(x)dx.$$

Simpson's method was used in the computation. From this value, the integral of $P(x)$ may be computed by the relation

$${}_0\int^1 P(x)dx = {}_0\int^1 Q(x)dx + \frac{1}{\alpha} + \frac{e^z}{\beta}.$$

A similar method was used to compute the value of

$${}_0\int^1 xP(x)dx.$$

III. *Replacement of integrals by summation based on a discrete gene frequency model:* In an actual population of N diploid individuals, gene frequency x takes on the discrete values, $0, \frac{1}{2N}, \ldots, 1 - \frac{1}{2N}, 1$.

If we denote by f_i the probability that $x = i/(2N)$, then $\phi(x)dx$ gives a good approximation to f_i if we substitute $1/(2N)$ for dx. This approximation is valid for unfixed classes, i.e. for $i = 1, 2, \ldots, 2N-1$. The probabilities of terminal classes may be obtained from

$$f_0 = \frac{1}{\alpha} P(\frac{1}{2N}) \frac{1}{2N}$$

$$f_{2N} = \frac{1}{\beta} P(1 - \frac{1}{2N}) \frac{1}{2N}$$

where $\alpha = 4Nu$ and $\beta = 4Nv$ (cf. WRIGHT 1931).

Therefore, (A.3) may be replaced by

$$\bar{L} = \frac{s[\sum_{i=1}^{2N-1} (\frac{i}{2N}) P(\frac{i}{2N}) + \frac{1}{\beta} P(1 - \frac{1}{2N})]}{\frac{1}{\alpha} P(\frac{1}{2N}) + \sum_{i=1}^{2N-1} P(\frac{i}{2N}) + \frac{1}{\beta} P(1 - \frac{1}{2N})}$$

where $P(x)$ is given by (A.2). A similar expression can be obtained readily for $\overline{L^2}$ from which V_L is obtained by (A.6).

Some numerical examples showing the agreement of the three methods are given in Table 2.

TABLE 2

Comparison of the mutation load by three methods of calculation, for $s=.02$, $h=.5$, $u=v=10^{-5}$

N	Method I	Method II	Method III
10	8.0278×10^{-3}	8.0278×10^{-3}	8.2203×10^{-3}
100	3.8741×10^{-4}	3.8738×10^{-4}	4.0167×10^{-4}
1000	2.0506×10^{-5}	2.0281×10^{-5}	2.0273×10^{-5}

LITERATURE CITED

CROW, J. F., 1958 Some possibilities for measuring selection intensities in man. Human Biol. **30**: 1–13.

FISHER, R. A., 1930, 1958 *The Genetical Theory of Natural Selection.* 2nd edition. Dover Publ., New York.

HALDANE, J. B. S., 1937 The effect of variation on fitness. Am. Naturalist **71**: 337–349.

KIMURA, M., 1961 Some calculations on the mutational load. Jap. Jour. Genet. **36** (Suppl.): 179–190.

Li, C. C., 1963 Decrease of population fitness upon inbreeding. Proc. Natl. Acad. Sci. U.S. **49**: 439–445.

MORTON, N. E., J. F. CROW, and H. J. MULLER, 1956 An estimate of the mutational damage in man from data on consanguineous marriages. Proc. Natl. Acad. Sci. U.S. **42**: 855–863.

MULLER, H. J., 1950 Our load of mutations. Am. J. Human Genet. **2**: 111–176.

ROBERTSON, A., 1962 Selection for heterozygotes in small populations. Genetics **47**: 1291–1300.

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

1937 The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci. U.S. **23**: 307–320.