EFFECTS OF RESTRICTED POPULATION SIZE AND INCREASE IN MUTATION RATE ON THE GENETIC VARIATION OF QUANTITATIVE CHARACTERS[1]

MASATOSHI NEI AND YOKO IMAIZUMI

Division of Genetics, National Institute of Radiological Sciences, Chiba, Japan

THE genetic variation of quantitative characters in an equilibrium population depends on the size of the population, and it is expected to be larger in a large population than in a small population. Introducing the concept of gene frequency distribution, WRIGHT (1931) was able to obtain a simple formula for the relationship between genetic variation and population size when the genes controlling quantitative characters are maintained by reversible mutations or a special kind of genic selection. Recently, ROBERTSON (1962) studied an important case of heterozygote advantage in terms of the frequency of heterozygotes and found that the genetic variation in small populations is highly dependent on the equilibrium gene frequency in an infinitely large population. Much more study is needed, however, in order to clarify further the general relationship between genetic variation and population size.

The mechanism through which the genetic variation of quantitative characters is maintained in natural populations is not well understood at present. There are four major hypotheses for the mechanism, i.e. (1) reversible mutation, (2) balance between mutation and selection, (3) overdominance or heterozygote advantage, and (4) optimizing selection. So far none of these hypotheses have been subjected to critical experimental test. If, however, these hypotheses lead to different relationships between the amount of genetic variation and the population size, then it would be possible experimentally to determine which mechanism is most important. We, therefore, examined the relationship theoretically with special reference to the first three hypotheses. There are several possible types of optimizing selection and this generally creates a difficulty in the mathematical treatment, especially such a model as proposed by KIMURA (1965). Note, however, that if genes are maintained by overdominance and their effects on a quantitative character are additive, an optimizing selection results (ROBERTSON 1956).

For the experimental test of the hypotheses it would also be helpful to know the relationship between the population size and the mean value of the character or the changes in mean value and genetic variation due to an increase in mutation rate. Mutation rate can easily be increased experimentally by the use of

radiation or other mutagenic agencies. Further, the relation between the population size and the changes in mean and variance would give useful information concerning the genetic effect of radiation on quantitative characters in man or other organisms.

MATHEMATICAL FORMULATION

Mean and genetic variance: Let us assume that a large number of loci, each with two alleles, are responsible for a quantitative character and that they act independently of each other both for the quantitative character and fitness. Thus, there are three possible genotypes for a locus, i.e. AA, Aa, and aa, where A and a refer to the alleles with positive and negative effects on the quantiative character. We denote the relative fitnesses and effects on the quantitative character of the three genotypes as follows:

Genotype	AA	Aa	aa
Frequency	q^2	$2q(1-q)$	$(1-q)^2$
Fitness	1	$1-h$	$1-s$
Quantitative character	a	d	$-a$

Here q represents the gene frequency of A. The contributions of this locus to the mean and additive genetic and dominance variances of the quantitative character are given by

Mean	$\gamma = (2q-1)a + 2q(1-q)d$
Additive variance	$v_a = 2q(1-q)[a + (1-2q)d]^2$
Dominance variance	$v_d = 4q^2(1-q)^2d^2$

We now assume that the values of h, s, a and d are the same for all relevant loci. The probability distribution of gene frequency in an equilibrium population of finite size is then given by

$$\phi(q) = Ce^{-4Nhq(1-q)-2Ns(1-q)^2} q^{4Nu-1} (1-q)^{4Nv-1} \tag{1}$$

where N is the effective population size, and u and v are the mutation rates of $a \to A$ and $A \to a$, respectively (WRIGHT 1937). Thus, the mean of the quantitative character is given by

$$Y = G + n \int_0^1 \gamma \phi(q)dq$$
$$= G + n [-a + 2(a+d)\mu'_1 - 2d\mu'_2] \tag{2}$$

where G and n stand for the background effect and the number of loci concerned, respectively, and μ'_1 and μ'_2 are the first and second moments of the gene frequency distribution, namely

$$\mu'_r = \int_0^1 q^r \phi(q)dq \tag{3}$$

In the same way the additive and dominance variances become

$$V_A = 2n[(a+d)^2\mu'_1 - (a+d)(a+5d)\mu'_2 + 4d(a+2d)\mu'_3 - 4d^2\mu'_4] \tag{4}$$

$$V_D = 4nd^2(\mu'_2 - 2\mu'_3 + \mu'_4) \tag{5}$$

respectively.

It is generally difficult to obtain the moments of gene frequency distribution analytically. However, in the case of neutral genes, which are maintained by reversible mutations, the moments are easily obtained and given by the beta functions as follows:

$$\mu'_r = \frac{B(U+r, V)}{B(U, V)} \tag{6}$$

where $U = 4Nu$ and $V = 4Nv$. Thus, for example, the additive variance in the case of no dominance $(d = 0)$ becomes

$$V_A = \frac{8nuva^2N}{(u+v)(U+V+1)}$$

which is equivalent to the formula obtained by WRIGHT (1931). In the case of genic selection $(h = \frac{1}{2}s)$ the moments are given by

$$\mu'_r = \frac{\Gamma(U+V)\Gamma(U+r)_1F_1(V, U+V+r, -2Ns)}{\Gamma(U)\Gamma(U+V+r)_1F_1(V, U+V, +2Ns)} \tag{7}$$

where $_1F_1$ denotes the confluent hypergeometric function.

Skewness and kurtosis: The skewness and kurtosis of a quantitative character are obtained in the same way and become

$$\gamma_1 = \frac{M_3}{M_2^{3/2}} = -\frac{\mu'_1 - 3\mu'_2 - 2\mu'_3}{(2n)^{1/2}(\mu'_1 - \mu'_2)^{3/2}} \tag{8}$$

$$\gamma_2 = \frac{M_4}{M_2^2} = 3 + \frac{1 - 6(\mu'_1 - \mu'_2)}{2n(\mu'_1 - \mu'_2)} \tag{9}$$

respectively.

When the genetic variation is maintained by reversible mutations, formulas (8) and (9) reduce to

$$\gamma_1 = \sqrt{\frac{(u+v)(u+v+1/(4N))}{2nuv}} \frac{u-v}{u+v+2/(4N)} \tag{10}$$

$$\gamma_2 = 3 + \frac{(u+v)(u+v+1/(4N)) - 6uv}{2nuv} \tag{11}$$

respectively. Formula (11) shows that in small populations the distribution of a quantitative character becomes leptokurtic unless n is very large. Further, if u differs extremely from v, the distribution shows a considerable skewness in large populations.

The skewness and kurtosis when there is dominance $(d > 0)$ or when the genes are subject to natural selection can be obtained in the same way or numerically, and it can be shown that the distribution is again leptokurtic when N is small.

Genetic load associated with quantitative characters: If the genes controlling a quantitative character are maintained by selection, they necessarily create a genetic load to the population. The magnitude of this genetic load can be obtained

by the same method as that used by KIMURA, MARUYAMA and CROW (1963), who studied the genetic load arising from deleterious mutations. For example, the segregational genetic load for an overdominant loci is given by

$$L_s = -h + s(1 - \mu'_2) - 2(s - h)(\mu'_1 - \mu'_2) \tag{12}$$

If Ns is small ($2Ns < 5$), this can be approximated by

$$L_s = \frac{(s - h) - h\,(u/v)\,e^{2Ns}}{1 + (u/v)\,e^{2Ns}} \tag{13}$$

Effects of an increase in mutation rate: When the mutation rate is increased artificially in a generation, the mean, genetic variance, and genetic load in the next generation may change through the change of gene frequency. If the mutation rate from a to A increases by m_1 and from A to a by m_2 in a generation, then the gene frequency of A increases as follows:

$$\Delta q = m_1 - (m_1 + m_2)q$$

If $m_1 = m_2 = m$, as is generally believed with the bristle characters of *Drosophila melanogaster* (cf. TOBARI and NEI 1965), Δq becomes $(1 - 2q)m$. Neglecting the effect of the higher orders of Δq, the change of mean is given by

$$\Delta Y = \int_0^1 \frac{dy}{dq} \Delta q \phi(q) dq$$

$$\tag{14}$$

$$= 2n\left[(a + d)m_1 - \{2dm_1 + (a + d)(m_1 + m_2)\}\,\mu'_1 + 2d(m_1 + m_2)\mu'_2\right]$$

In the same way the changes in genetic variance and genetic load can be obtained.

NUMERICAL COMPUTATIONS

Methods of computation: As shown in the preceding section, the mean, genetic variance or other quantities in a population of finite size can be expressed as a function of the moments of the gene frequency distribution. The evaluation of the moments, however, usually necessitates a numerical integration, since they are not generally expressible in terms of elementary functions. If the right side of (1) with C omitted is denoted by $P(q)$, the moments are obtained by

$$\mu'_r = \frac{\int_0^1 q^r P(q) dq}{\int_0^1 P(q) dq} \tag{15}$$

In the computation of the values of integrals in the above formula we used Simpson's method. When the population size is small, the value of $P(q)$ may become $+\infty$ both at $q = 0$ and 1, so that Simpson's method is not directly applicable. In this case we used the method devised by KIMURA (1963). In some cases the integrals in (15) can be expressed by elementary functions and these serve as checks on the numerical integration: (i) When the effective population size is small, the first to fourth moments are almost identical and given by

$$\mu' = \frac{1}{1 + (v/u)\,e^{-2Ns}} \tag{16}$$

approximately. (ii) When the effective population size is extremely large, μ'_r is given by $\hat{q}^r$. (iii) In the case of neutral genes the moments are obtained directly from (6). (iv) In the special case of $U = V = 1$ and $h = \frac{1}{2}$, a direct integration is possible; in this case the moments are also obtainable from (7). (v) In the case of $U = V = 1$ and $h = 0$,

$$\mu'_1 = \frac{C}{4Ns}\left(1 - e^{-2Ns}\right)$$

$$\mu'_3 = \frac{C}{4Ns}\left[\frac{1}{2Ns} - e^{-2Ns}\left(1 + \frac{1}{2Ns}\right)\right]$$

where C is $\sqrt{2Ns/\pi}/erf(2\sqrt{Ns})$, $erf(x)$ being $\int_0^x 1/\sqrt{2\pi}\, e^{-t^2/2}dt$.

In the numerical computation we chose the values of s and h so as to give three different levels of the equilibrium gene frequency at $N = \infty$, namely 0.5, 0.7, and 0.9 (see Table 1). Note that, when s and h are small, the equilibrium gene frequency at $N = \infty$ is affected by both the forward and backward mutations. In all cases $u = 10^{-5}$ was used, while for the reverse mutation $v = (3/7) \times 10^{-5}$ was used in the cases of (4), (14), and (15), and $v = (1/9) \times 10^{-5}$ in the cases of (9), (16), and (17). In all other cases we assigned 10^{-5} to v. It may be noted that cases (1), (4), and (9) represent neutral genes, and cases (8), (12), and (13) genes maintained by a balance between mutation and selection. The other cases all relate to genes showing heterozygote advantage or overdominance. The selection coefficients in (14), (15), (16), and (17) are almost equal to those in (5), (6), (10), and (11) respectively, though the mutation rates are not the same. The latter group of cases was set up just to see the effect of spontaneous mutation rate in small populations.
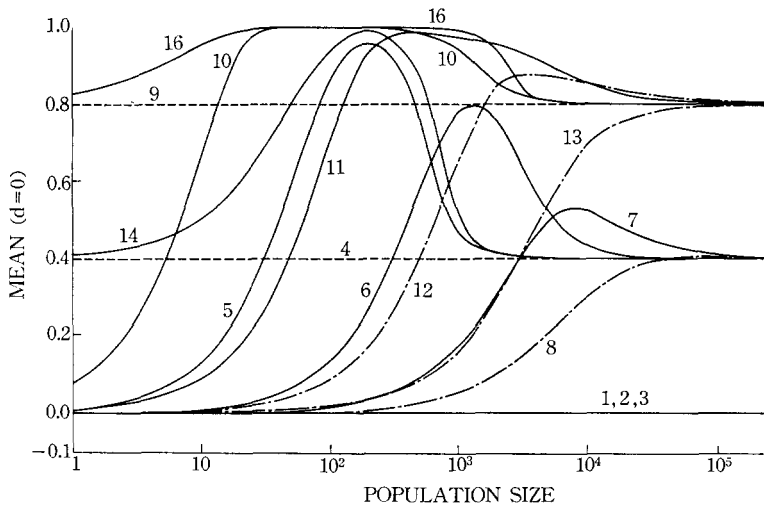

FIGURE 1.—Relations between the mean of a quantiative character and population size $(d = 0)$. The broken lines refer to no selection and the solid lines to overdominant selection. The lines in dot and dash refer to genic or recessive selection. The numbers given to each curve represents the case number in Table 1.

Mean: If there is no dominance for quantitative characters or $d = 0$, expression (2) reduces to $Y = G + na(2\mu'_1 - 1)$. Thus, the relation between the mean and population size is determined by $2\mu'_1 - 1$. The values of $2\mu'_1 - 1$ are given in Figure 1. For cases (1), (2), and (3) $\mu'_1 = 0.5$ for all values of N, so that the mean of a quantitative character is independent of population size. Also, in cases (4) and (9) the means are independent of population size, though the values are not the same as those of the first three cases. Case (5) shows that the mean value first increases logistically with an increase of N, reaching the maximum value at about $N = 200$, and then begins to decrease, approaching the value of $N = \infty$. The large values of the means in populations of the order of 200 has been brought about by the mean gene frequencies drifting away from the equilibrium values. That the gene frequencies in small populations are deviated far from the equilibrium values when the equilibrium gene frequencies at $N = \infty$ are not 0.5 has been noted by ROBERTSON (1962). It is seen from (6) and (7) that a similar pattern of change in mean obtains also for the smaller values of s and h though the pattern is less conspicuous. Comparison of (5) with (14) or (10) with (16) shows that, if $u > v$, the mean value is further increased in populations of intermediate size. It is also observed that the mean in small populations is affected more strongly by mutation than by selection intensity.

In the case of complete dominance (cases (8) and (12)), the mean value in populations of intermediate size is also slightly larger than the value at $N = \infty$. When there is selection without dominance or genic selection (13), the relation between the mean and population size is of S shape.

When the effects of genes on quantitative characters show complete dominance $(d = a)$, expression (2) becomes $Y = G - na(1 - 4\mu'_1 + 2\mu'_2)$. The values of $-1 + 4\mu'_1 - 2\mu'_2$ are given in Figure 2. The mean value is no longer independent of the population size even in the case of neutral genes (1, 4, and 9). The same
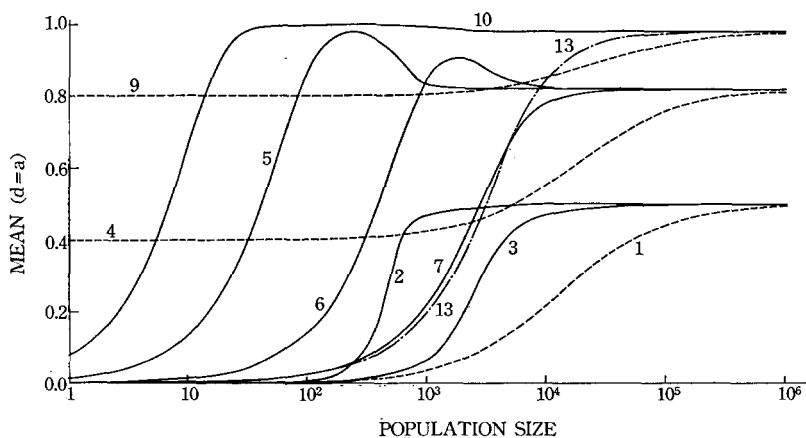

FIGURE 2.—Relations between the mean of a quantitative character and population size $(d = a)$. The broken lines refer to no selection and the solid lines to overdominant selection. The line in dot and dash refers to genic selection. The number given to each curve represents the case number in Table 1.

is true for the cases of (2) and (3), i.e. overdominance. Other cases show the same type of change in mean as in the case of $d = 0$, but the difference between the maximum value and the value at $N = \infty$ is smaller.

Theoretically, it is possible to make d equal to $-a$. In this case $Y = G - na(1 - 2\mu'_2)$. Thus, the relation between Y and N is somewhat different from that of the case where $d = a$. However, we will not consider this any more, since it appears important only in limited cases.

Additive genetic and dominance variances: With $d = 0$ the additive genetic variance becomes $2na^2(\mu'_1 - \mu'_2)$. The values of $\mu'_1 - \mu'_2$ are given in Figure 3. The relation between the additive genetic variance and effective size is rather simple, and in all cases an S curve obtains. Selection, whether genic, recessive or overdominant, generally increases the variance in small populations. Thus, when the fitness or heterozygotes is higher than those of both homozygotes by 1% (case 2), the genetic variance in a population of the order of 1000 is almost equal to that in a population of more than 100,000 in the case of neutral genes. The same is true also for the case of (5), when compared with case (4). The genetic variability maintained in a small population is generally larger when genes are subject to strong selection than when subject to weak selection. It was also found that the genic selection with small values of s and h (13) has a stronger effect on the genetic variance than an overdominant selection with rather large absolute values of s and h (11).

In a study of the effect of overdominance on the frequency of heterozygotes, ROBERTSON (1962) reached the conclusion that, while it can conserve a larger genetic variability than in the case of neutral genes in a region of $0.15 \sim 0.85$ of
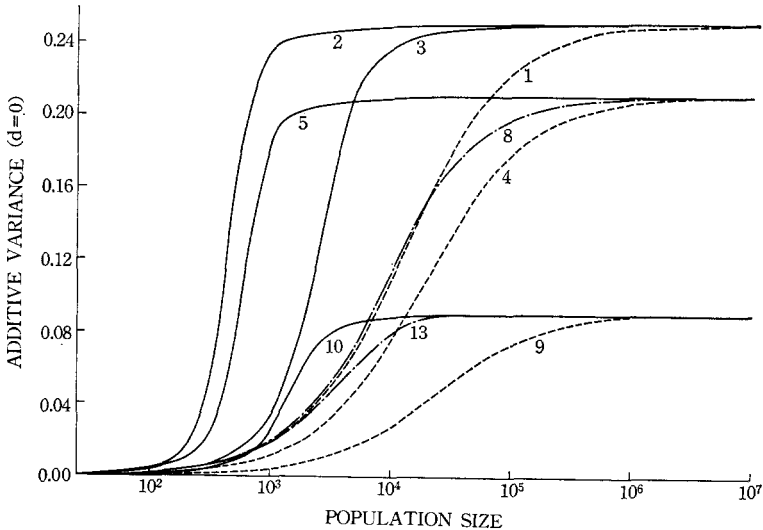

FIGURE 3.—Relations between the additive genetic variance and population size $(d = 0)$. The broken lines refer to no selection and the solid lines to overdominant selection. The lines in dot and dash refer to genic or recessive selection. The number given to each curve refers to the case number in Table 1.
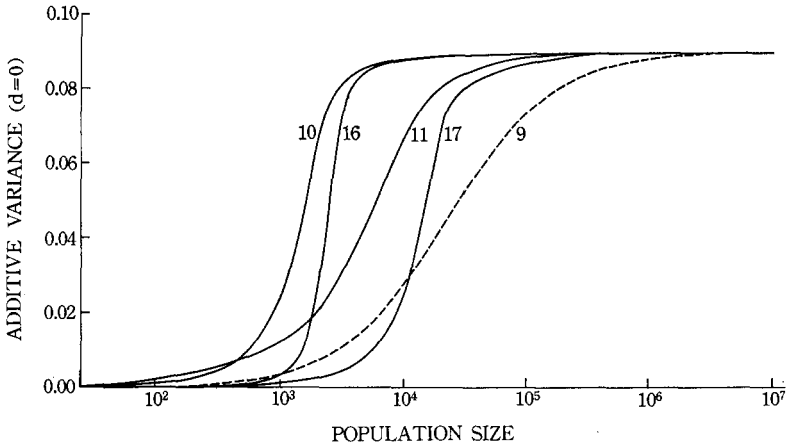
FIGURE 4.—Effect of overdominance in fitness on the additive genetic variance when the equilibrium gene frequency is 0.9. The broken line refers to no selection and the solid lines to overdominant selection. The number given to each curve represents the case number in Table 1. It is seen that curves 16 and 17 cross 9.

the equilibrium gene frequency at $N = \infty$, it may reduce the genetic variability outside the region if $N(s - 2h)$ is smaller than a certain critical value. This conclusion is based on a study of rather small populations with equal mutation rates of u and v. However, if the population size is very large, the genetic variance under selection again becomes smaller than that in the case of no selection, as seen
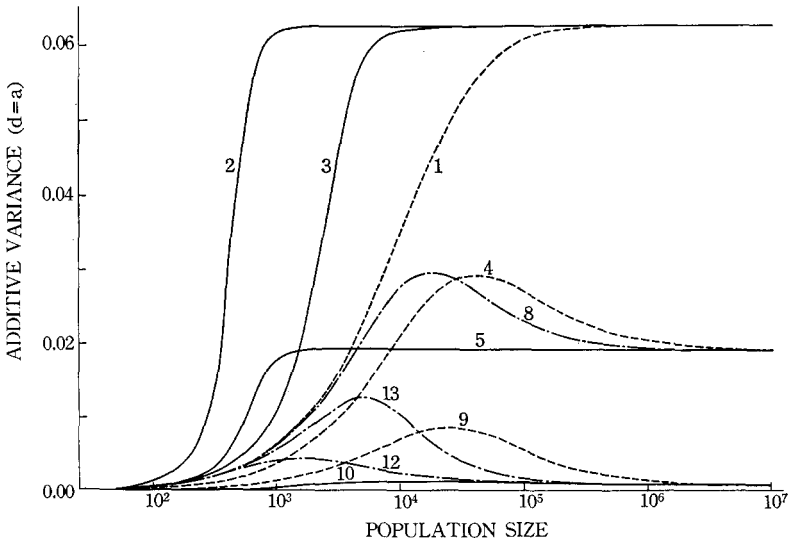

FIGURE 5.—Relations between the additive genetic variance and population size $(d = a)$. The broken lines refer to no selection and the solid lines to overdominant selection. The lines in dot and dash refer to genic or recessive selection. The number given to each curve represents the case number in Table 1.

from the comparison of curves (1) and (10) in Figure 3. If, on the other hand, the mutation rates for neutral genes are assigned so as to give the same equilibrium gene frequency as that specified by a given type of selection, the genetic variance in large populations in larger for non-neutral genes than for neutral genes [(9) vs. (10)]. In this case, however, the ROBERTSON effect disappears. In order for this effect to hold, the mutation rates for non-neutral genes should be the same as or close to those for neutral genes, as seen from Figure 4. In the case of (16) in Figure 4 the ROBERTSON effect arises for $N < 10^3$ and in the case of (17) for $N < 10^4$. No such effects were observed for the cases of (14) and (15).

When $d = a$, the relationship between the additive genetic variance and population size is slightly changed, as seen from Figure 5. With the equilibrium gene frequency of 0.7 or 0.9 and weak or no selection, the genetic variance in a population of intermediate size, i.e. around 10^4, is larger than the value in an infinitely large population. This is caused by the fact that, when there is weak or no selection, the gene frequencies are affected by random genetic drift more severely than when there is strong selection. It is of interest to see that, when the equilibrium gene frequency is 0.9 and N is smaller than 10^4, genic selection (13) is capable of maintaining a larger genetic variance than no selection (9). Overdominant selection (10), however, maintains only a small variance. This is due to the fact that under strong selection the moments of gene frequency quickly reach the values of $N = \infty$ as N increases. The same is true for the case of (5).

The numerical values of dominance variance, i.e. $V_D/4nd^2 = \mu'_2 - 2\mu'_3 + \mu'_4$, are graphically shown in Figure 6. The relationship between the dominance
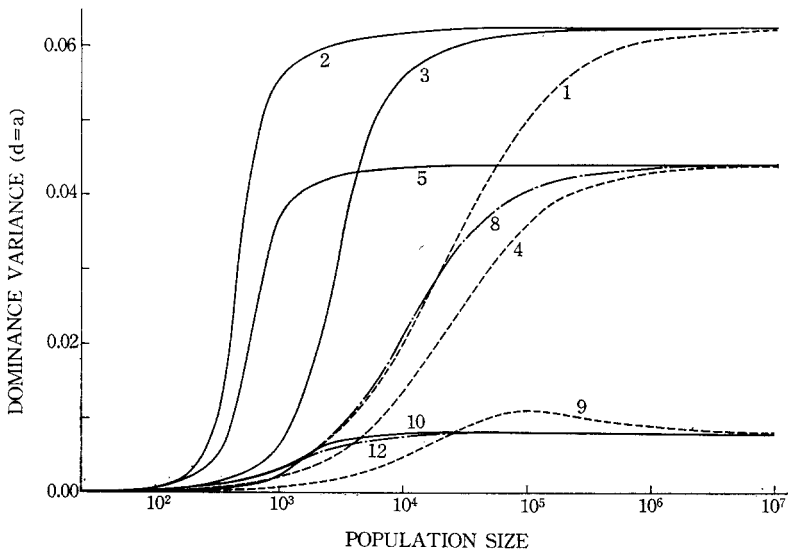

FIGURE 6.—Relations between the dominance variance and population size $(d = a)$. The broken lines refer to no selection and the solid lines to overdominant selection. The lines in dot and dash refer to genic or recessive selection. The number given to each curve represents the case number in Table 1.

variance and population size is almost the same as that between the additive genetic variance and population size, except that with some sets of small selection coefficients (e.g. 12) the intermediate size of N no longer shows the largest value.

Genetic load: The genetic loads for representative cases are plotted against the population size in Figure 7. The relation between the mutation load and population size has been fully investigated by KIMURA, MARUYAMA, and CROW (1963), and our results on this type of load are almost the same as theirs. One exception is that the genetic load created by small selection coefficients becomes smaller than the mutation rate when N is large. For example, in the case of (8) the genetic load at $N = \infty$ becomes 0.6×10^{-5} for a locus. This smaller load than the mutation rate is due to the fact that with weak selection the reverse mutation $a \rightarrow A$ is no longer negligible when $u = v$. As indicated by KIMURA *et al.* (1963), the mutation load increases as N decreases. Thus, in the case of (12) the genetic load for $N = 100$ is almost 40 times the load for $N = \infty$ (Figure 7). If there are 100 such loci controlling a quantitative character, the reduction in fitness is $1 - e^{-100 \times 0.9 \times 10^{-5}} = 0.09\%$ for $N = \infty$ and 3.5% for $N = 100$.

The relation of the segregation load to the population size is somewhat different from that of the mutation load. First, the difference in the segregation load between large and small populations is not generally so great as the difference in the mutation load. In general, the load for segregational loci is quite high even when $N = \infty$ but does not increase so strikingly with a decrease of population size as does the load for mutational loci. In the case of (2), for example, the total load for 100 loci is $100 \times 5 \times 10^{-3} = 0.5$ when $N = \infty$ and 1 when $N = 100$. The relative fitnesses of the two populations are 0.61 and 0.43 respectively. Second,
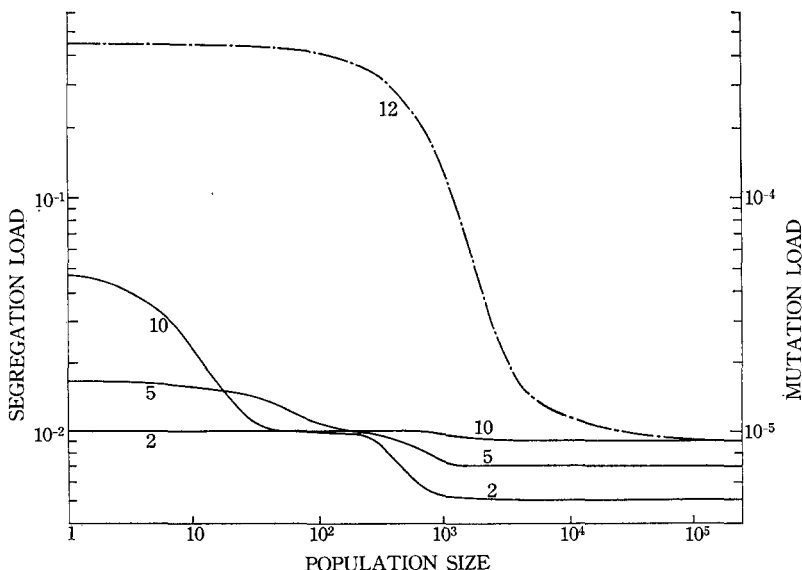

FIGURE 7.—Relations between the genetic load and population size. Curve 12 refers to the mutation load and the others to the segregation load (cf. Table 1).

with some sets of selection coefficients the genetic load shows a two-step change as N increases, as in the cases of (5) and (10).

Changes of mean due to an increase in mutation rate: When $d = 0$ and $m_1 = m_2 = m$, the amount of change in mean is given by $2nma(1 - 2\mu'_1)$. The dependence of this change on the population size can be seen from Figure 1, where the value of $2\mu'_1 - 1$ is given. The interpretation of this figure has already been discussed. In the case of complete dominance $(d = a)$, $\Delta Y = 4nma(1 - 3\mu'_1 + 2\mu'_2)$. Numerical computations have shown that the relation of ΔY to N is roughly the reverse of the relation of Y to N and ΔY is generally larger in small populations than in large populations.

Changes of genetic variances: The changes of additive variance due to an increase in mutation rate when $d = 0$ are shown in Figure 8. The relation of ΔV_A to N is again roughly the reverse of the relation of V_A to N. In small populations most loci are fixed, so that almost every mutation creates a new genetic variation. A large population, on the other hand, contains both positive and negative alleles, so that the frequency of heterozygotes is hardly affected, unless the equilibrium gene frequency deviates substantially from 0.5. If the equilibrium gene frequency is 0.7, ΔV_A at $N = \infty$ is less than one fifth of the value at $N < 100$.

The amounts of changes in additive genetic variance with $d = a$ are shown in Figure 9. When the equilibrium gene frequency is 0.5, ΔV_A decreases monotonically with an increase of N, but if it differs from 0.5, this value decreases far below the level of $N = \infty$ in populations of intermediate size. In an extreme case of (10) the value of ΔV_A is almost zero for $N = 10^2 \sim 10^3$. Apparently, these decreases are caused by the gene frequency drifting away from the equilibrium value.

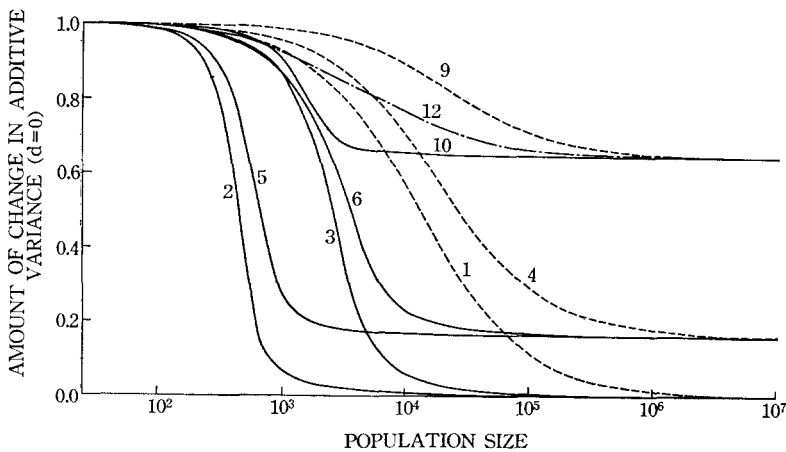The change in dominance variance in a population of $N < 100$ is negligibly

FIGURE 8.—Amounts of changes in additive genetic variance due to increase in mutation rate in populations of finite size $(d = 0)$. The broken lines refer to no selection and the solid lines to overdominant selection. The line in dot and dash refers to recessive selection. The number given to each curve represents the case number in Table 1.
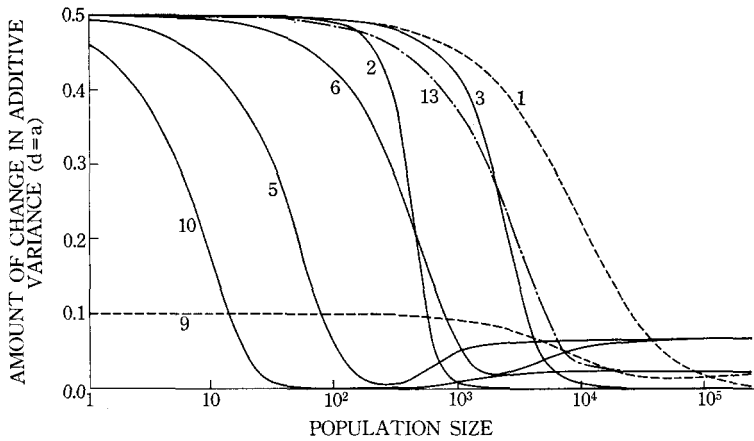
FIGURE 9.—Amounts of changes in additive genetic variance due to increase in mutation rate in populations of finite size $(d = a)$. The broken lines refer to no selection and the solid lines to overdominant selection. The line in dot and dash refers to genic selection. The number given to each curve represents the case number in Table 1.

small but it gradually increases as N increases (Figure 10). When the equilibrium gene frequency is 0.9, ΔV_D monotonically increases to reach the largest value of $N = \infty$. However, if it is 0.5, ΔV_D again decreases and becomes 0 at $N = \infty$ (1, 2, and 3). This is understandable if we note that the change in dominance variance can be written as $8\Sigma mq(1 - q)(1 - 2q)^2d^2$, where Σ indicates the summation over all relevant loci. This quantity is 0 if q is any of 0, 0.5, and 1. In small populations q is mostly 0 or 1, whereas in large populations q is 0.5 at
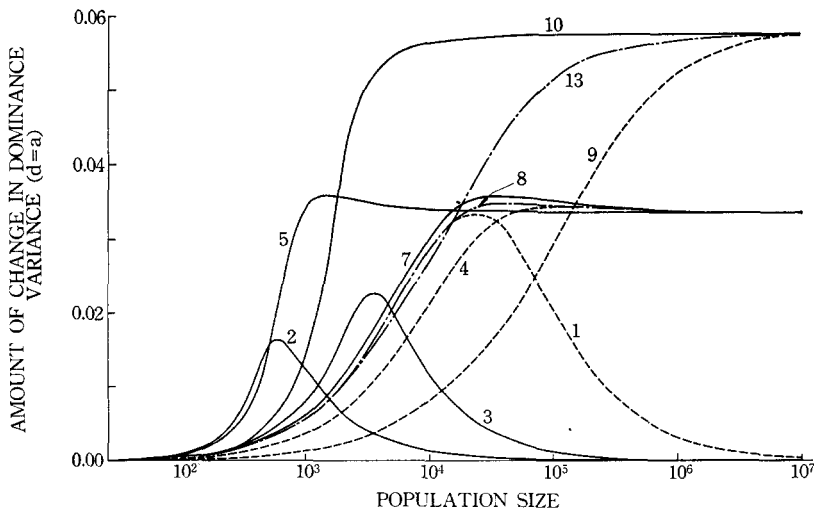

FIGURE 10.—Amounts of changes in dominance variance due to increase in mutation rate in populations of finite size $(d = a)$. The broken lines refer to no selection and the solid lines to overdominant selection. The line in dot and dash refers to genic selection. The number given to each curve represents the case number in Table 1.
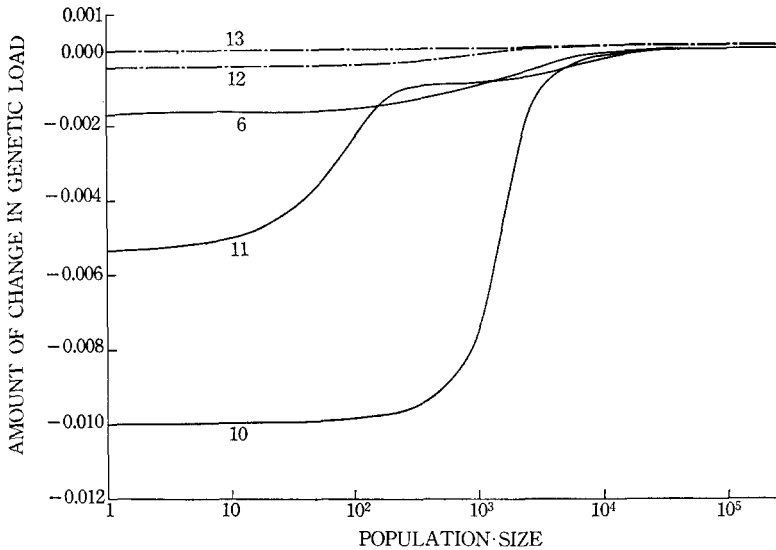
FIGURE 11.—Amounts of changes in genetic load due to increase in mutation rate. The solid lines refer to segregation load and the lines in dot and dash to mutation load. The number given to each curve represents the case number in Table 1.

most loci. In both cases, therefore, ΔV_D is nearly 0. In populations of intermediate size, however, q is flatly distributed with the mean value of 0.5, and those loci at which $0 < q < 0.5$ or $0.5 < q < 1$ contribute to the value of ΔV_D.

Changes of genetic load: When the genetic variation is maintained by mutation and selection, the amount of change in genetic load is generally small, as seen from Figure 11 (of course, if s and h are large, this change could be large). In an infinitely large population the change in genetic load is positive as it should be, but in a small population it can be negative if dominance for fitness exists, as in the case of (12). In this case such mutations as $AA \rightarrow Aa$ show no effect on fitness, while the mutations such as $aa \rightarrow Aa$ increase fitness. Therefore, the genetic load decreases.

In the case of overdominance the genetic load is almost independent of mutation rate if N is large. However, if N is small, the genetic load is reduced by increase in mutation rate (Figure 11). This reduction in the genetic load or the increase of fitness is caused by the increase of heterozygote frequency due to mutations.

<center>DISCUSSION</center>

From the results given above it is obvious that the magnitudes of genetic parameters such as mean, variance and genetic load depend on the population size as well as the mechanism through which the genetic variation is maintained. These results may be useful in discriminating between different mechanisms of maintaining the genetic variation. For example, if the mean value is independent of population size and the genetic variance in populations of less than 1000 is

very small compared with that of a large population (of course, assuming that the mutation rate is of the order of 10^{-5}), it may be conjectured that the genetic variation is maintained by reversible mutations or very slight overdominant selection. When overdominant seelction operates and the equilibrium gene frequencies are far removed from 0.5, the ROBERTSON effect takes place, and the genetic variance in small populations could be smaller than in the case of reversible mutations. In this case, however, the mean value changes with change of population size and the genetic load is considerably larger even in large populations. The genetic load associated with a quantitative character may be measured by the method developed by HALDANE (1954) in practice, assuming no interaction between genetic and environmental effects. When the genetic variation is maintained by the balance between mutation and selection, the relation between the genetic variance and population size may again look like the relation for neutral genes. In this case, however, the genetic load in small populations would be much larger than that in large populations. For the purpose of discriminating between the different mechanisms of maintaining the genetic variation information on the changes in genetic parameters may also be useful.

If there is dominance for quantitative characters, the relations of genetic parameters with population size are somewhat different from those in the case of no dominance, but the relations may again be used for discriminating between the different mechanisms. The degree of dominance of quantitative characters can be studied by the comparison of full sib and parent-offspring correlations or other procedures specifically designed for the purpose.

Quantitative characters are often subject to a centripetal selection. Several models of optimizing selection have been put forward to explain this type of selection, though most of the models have been proved to be inadequate for maintaining the genetic variation permanently (FISHER 1930; WRIGHT 1935; KOJIMA 1959; LEWONTIN 1964). Recently, KIMURA (1965) proposed a new model based on a number of assumptions, of which the adequacy is not necessarily demonstrated.

On the other hand, ROBERTSON (1956) has shown that, if the individual genes are maintained by overdominance, a centripetal selection results. This was extended to the case of balance between mutation and selection by KIMURA (1958). Thus, if any of these kinds of selection leads to the same type of centripetal selection as observed in nature or in experiments, optimum models may be unnecessary. For the overdominance or mutation-selection model to be correct, several properties other than those considered here should be satisfied (cf. ROBERTSON 1959; KIMURA 1960).

In general, if the mutation-selection model is correct, the gene frequency in large populations is usually far from 0.5, so that the genetic variance per locus is very small. Consequently, a much larger number of loci should be concerned, compared with the case of overdominance, in order to have a given amount of genetic variability. Further, if the average gene frequency differs greatly from 0.5, the long term responses to the upward and downward artificial selections should be drastically different, provided that, as is likely, the effects of genes on

the quantitative character and fitness are correlated. It is also expected that the response to selection is very small in early generations and gradually increases as generation proceeds, since it depends on the quantity $q(1-q)$. However, the selection experiments conducted with Drosophila or mice do not support any of these predictions. On the other hand, the overdominant selection usually creates a large genetic load in large populations compared with the mutation-selection scheme, and this is one of the reasons why some authors object to the overdominance hypothesis. If, however, the heterozygote advantage is only of 1% or less, the genetic load created by some 100 such loci would not be so large. In small populations the genetic load is quite large both for mutational and segregational loci.

Natural populations are not necessarily a single random mating unit but usually subdivided into many subpopulations in which random mating takes place. In such populations the gene frequency distribution for neutral genes is given by

$$\phi(q) = \frac{q^{U'-1}(1-q)^{V'-1}}{B(U', V')}$$

where

$$U' = \frac{4Nv\,[m(1-r)+u+v]}{u+v}$$

$$V' = \frac{4Nv\,[m(1-r)+u+v]}{u+v}$$

and m and r stand for the migration rate and correlation of gene frequency between neighboring subpopulations (cf. KIMURA and WEISS 1964), and N for
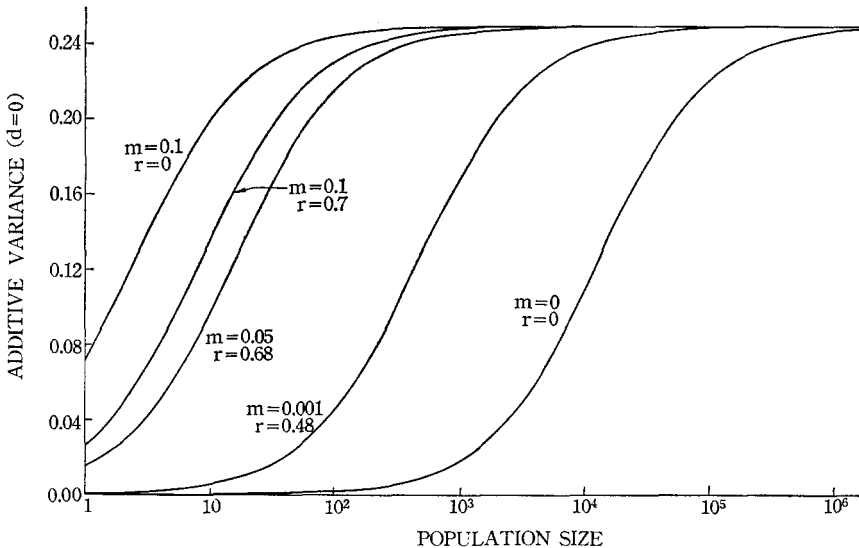

FIGURE 12.—Effect of migration on the additive genetic variance $(d=0)$ in 'island' and 'stepping stone' models. m and r represent the migration rate and correlation of gene frequency between neighboring subpopulations. The forward and backward mutation rates are each 10^{-5}.

the effective size of subpopulations. In WRIGHT's "island model" (WRIGHT 1943), r is always 0, while in the "stepping stone model" by KIMURA and WEISS (1964) r is a function of m. In the case of two dimensional distribution r is 0.70 for $m = 0.1$, 0.68 for $m = 0.05$, and 0.48 for $m = 0.001$. Using the above distribution, we can obtain the mean, genetic variance, etc. of quantitative characters as before. The result for the additive genetic variance with $d = 0$ is shown in Figure 12. In this computation $u = v = 10^{-5}$ was used.

It is clear that in the case of island model with $m = 0.1$ the additive genetic variance in subpopulations of $N = 100$ is almost the same as that in an infinitely large population. It is also noted that the genetic variance for $N = 100$ and $m = 0.1$ is of the same order of magnitude as that for $N = 10^6$ and $m = 0$. The genetic variance for stepping stone model is slightly smaller than that for island model when the migration rate is the same. It is, however, remarkable that the migration of only 0.1% can increase the genetic variance considerably compared with the case of complete isolation. NEI and IMAIZUMI (1966) have estimated the effective size and migration rate of the so-called isolated populations in Japan. The estimates obtained are 1993 and 0.06 respectively. It may, therefore, be inferred that the genetic variance of quantitative characters in these populations is hardly different from that in the general Japanese population.

Information on the changes of genetic parameters due to increases in mutation rate may be useful in the evaluation of genetic effect of radiation in man or other organisms. For example, if there is no dominance ($d = 0$) and no selection and the forward and backward mutations occur with equal frequency, the means of quantitative characters are hardly affected in any population. The results obtained with *Drosophila melanogaster* may reflect this circumstance (TOBARI and NEI 1965). On the other hand, if there is dominance or selection or the forward mutation rate far exceeds the backward mutation rate, the means could be lowered by radiation. One such case is known with the maze learning ability of rats (NEWCOMBE and McGREGOR 1964), though the underlying genetic mechanism is not clear. It is of interest to see that under the assumptions made in this paper the change in mean is larger in small populations than in large populations.

The effect of radiation on the variance may be evaluated from Figures 8 and 9, assuming the constancy of environmental variance. It is clear that the amount of increase of genetic variance is larger in small populations than in large populations. In such a case as (1) in Figure 8, the increase of genetic variance in populations of less than 1000 is considerable, while in populations of larger than 10^5 the genetic variance hardly changes. The degree of dominance and the selection scheme affect the relation between N and ΔV_A or $\Delta V_A + \Delta V_D$ but the general pattern remains the same. This indicates that the genetic hazard of radiation to man may not be the same for small-sized isolated populations and large general populations.

TOBARI and NEI (1965) examined the effect of X rays on the genetic variance of abdominal bristles of *D. melanogaster*, using an approximately equilibrium population of about 8,000 individuals, and obtained almost the same amount of

variance increase as observed with an isogenic population. Assuming the effective population size to be 4,000, this result suggests that the genetic variation is maintained largely by reversible mutations or very slight selection. Note that this character is known to show little dominance and mutations occur in both positive and negative directions with an almost equal frequency. Further, the upward and downward selections so far conducted suggest that the average gene frequency is not far from 0.5.

Finally, it should be noted that the present analysis is based on the assumptions of no epistasis and no multiple alleles. If these assumptions are not fulfilled, the relationship between the genetic variation and population size may be somewhat different, but the general pattern appears to be the same, unless epistatic effects are large.

SUMMARY

The effects of population size and increase in mutation rate on the mean and genetic variance of quantitative characters and the genetic load associated with these characters have been theoretically examined in relation to the mechanisms through which the genetic variation is maintained in the population. For this purpose three different mechanisms, i.e. reversible mutation, balance between mutation and selection, and overdominant selection (heterozygote advantage) are considered.—If a quantitative character is controlled by neutral genes with reversible mutation and their effects on the quantitative character are additive without dominance, the mean of the character is independent of the population size (N). If, however, there is dominance (more fit alleles are dominant), the mean is no longer independent of N. Under overdominant selection the mean first increases logistically with an increase of N, reaching the maximum value in populations of intermediate size, and then begins to decrease to approach the value of $N = \infty$, except the case where the homozygotes for the alleles with positive and negative effects are equally fit. Mutation-selection scheme shows a similar but less conspicuous relation between the mean and population size.— The genetic variance of quantitative characters increases logistically with an increase of N. When there is selection, the genetic variance in small populations is generally larger than that of the case of no selection. If, however, the equilibrium gene frequency at $N = \infty$ is outside of 0.15-0.85, overdominant selection may reduce the genetic variance in small populations more than the level of the case of no selection.—The genetic load associated with a quantitative character is greater in small populations than in large populations both for mutational and segregational loci. In large populations the genetic load is generally much larger for segregational loci than for mutational ones, but in small populations both the segregation and mutation loads become of the same order of magnitude if selection intensity is nearly equal. The amounts of changes in mean and genetic variance due to an increase in mutation rate are generally greater in

small populations than in large populations, though they depend on various factors such as the forward and backward mutation rates and the mechanism of maintaining the genetic variation.—The bearings of the results obtained on the experimental discrimination between the different mechanisms of maintaining the genetic variation and the evaluation of genetic hazard of radiation to man have been discussed. The effect of migration on the genetic variation has also been examined by using the "island" and "stepping stone" models.

LITERATURE CITED

FISHER, R. A., 1930 *The Genetical Theory of Natural Selection.* Clarendon Press, Oxford.

HALDANE, J. B. S., 1954 The measurement of natural selection. Proc. 9th Intern. Congr. Genet. : 480–487.

KIMURA, M., 1958 On the decrease of fitness due to deviation from optimum in quantitative characters. Ann. Rep. Nat. Inst. Genet. (Japan) **8**: 106–107. —— 190 Relative applicability of the classical and the balance hypotheses to man. Especially with respect to quantitative characters. J. Radiation Res. **1**: 155–164. —— 1963 Numerical methods used to evaluate the mean and the variance of the mutational load with genic selection. Genetics **48**: 1310–1312. —— 1965 A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proc. Natl. Acad. Sci. U.S. **54**: 731–736.

KIMURA, M., T. MARUYAMA, and J. F. CROW, 1963 The mutational load in small population. Genetics **48**: 1303–1312.

KIMURA, M., and G. H. WEISS, 1964 The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics **49**: 561–576.

KOJIMA, K., 1959 Stable equilibria for the optimum model. Proc. Natl. Acad. Sci. U.S. **45**: 989–993.

LEWONTIN, R. C., 1964 The interaction of selection and linkage. II. Optimum models. Genetics **50**: 757–782.

NEI, M., and Y. IMAIZUMI, 1966 Genetic structure of human populations. II. Differentiation of blood group gene frequencies among isolated populations. Heredity **21**: 183–190.

NEWCOMBE, H. B., and J. F. McGREGOR, 1964 Learning ability and physical well-being in offspring from rat populations irradiated over many generations. Genetics **50**: 1065–1081.

ROBERTSON, A., 1959 The effect of selection against extreme deviants based on deviation or on homozygosis. J. Genet. **54**: 236–248. —— 1962 Selection for heterozygotes in small populations. Genetics **47**: 1291–1300.

TOBARI, I., and M. NEI, 1965 Genetic effects of X rays on quantitative characters in a heterogeneous population of *Drosophila melanogaster.* Genetics **52**: 1007–1015.

WRIGHT, S., 1931 Evolution in mendelian populations. Genetics **16**: 97–159. —— 1935 The analysis of variance and the correlations between relatives with respect to deviations from an optimum. J. Genet. **30**: 243–256. —— 1937 The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci. U.S. **23**: 307–320. —— 1943 Isolation by distance. Genetics **28**: 114–138.

APPENDIX: CHANGE OF LINKAGE DISEQUILIBRIUM, OWING TO MUTATION

The departure from a random combination of gene frequencies may be measured by the so-called linkage disequilibrium. For two loci each with two alleles A, a and B, b this linkage disequilibrium is defined as

$$D = PS - QR$$

where P, Q, R, and S stands for the frequencies of gametes AB, Ab, aB and ab respectively. If the mutation rates of $a \rightarrow A$ and $b \rightarrow B$ are each m_1 and those of $A \rightarrow a$ and $B \rightarrow b$ are m_2, the transition matrix of gamete frequencies due to mutation become as given in Table 2. Thus, neglecting the second order terms of mutation rate, we have

$$\Delta P = -2m_2 P + m_1(Q+R)$$
$$\Delta Q = m_2(P-Q) - m_1(Q-S)$$
$$\Delta R = m_2(P-R) - m_1(R-S)$$
$$\Delta S = m_2(Q+R) - 2m_1 S$$

where ΔP, ΔQ, ΔR, and ΔS are the amounts of changes of P, Q, R, *and* S, respectively, due to mutation. The amount of change of D is then obtained as follows:

$$\Delta D = \frac{\partial D}{\partial P} \Delta P + \frac{\partial D}{\partial Q} \Delta Q + \frac{\partial D}{\partial R} \Delta R + \frac{\partial D}{\partial S} \Delta S$$

$$= -2(m_1 + m_2)D$$

This indicates that the change of linkage disequilibrium is always such as to reduce the pre-existing disequilibrium.

As seen from the above, mutation has an effect analogous to that of recombination. When there is no selection, therefore, we have

$$D_t = (1-r-2m_1-2m_2)^t D_0$$

where D_t is the linkage disequilibrium in the tth generation and r stands for the recombination value between the two loci. Hence, a continuous exposure to radiation would speed up the breakdown of linkage disequilibrium.

TABLE 1

Selection coefficients employed for numerical computations

Case number	Type of selection	s	h	$\hat{q}^*$ $N=0$	$N=\infty$
(1)	Neutral	0	0	0.5	0.5
(2)	Overdominant	0	—0.01	0.5	0.5
(3)	Overdominant	0	—0.001	0.5	0.5
(4)	Neutral	0	0	0.7	0.7
(5)	Overdominant	0.013397	—0.01	0.5	0.7
(6)	Overdominant	0.001397	—0.001	0.5	0.7
(7)	Overdominant	0.000197	—0.0001	0.5	0.7
(8)	Recessive	0.000063	0	0.5	0.7
(9)	Neutral	0	0	0.9	0.9
(10)	Overdominant	0.08089	—0.01	0.5	0.9
(11)	Overdominant	0.00889	—0.001	0.5	0.9
(12)	Recessive	0.000889	0	0.5	0.9
(13)	Genic	0.000178	0.000089	0.5	0.9
(14)	Overdominant	0.013333	—0.01	0.7	0.7
(15)	Overdominant	0.001333	—0.001	0.7	0.7
(16)	Overdominant	0.08032	—0.01	0.9	0.9
(17)	Overdominant	0.00832	—0.001	0.9	0.9

* Equilibrium gene frequencies at two extremes.

TABLE 2

Transition matrix of gamete frequencies due to mutation

	AB P	Ab Q	aB R	ab S
AB	$(1-m_2)^2$	$(1-m_2)m_1$	$m_1(1-m_2)$	m_1^2
Ab	$(1-m_2)m_2$	$(1-m_2)(1-m_1)$	m_1m_2	$m_1(1-m_1)$
aB	$m_2(1-m_2)$	m_2m_1	$(1-m_1)(1-m_2)$	$(1-m_1)m_1$
ab	m_2^2	$m_2(1-m_1)$	$(1-m_1)m_2$	$(1-m_1)^2$