BEHAVIOR OF NEUTRAL MUTANTS INFLUENCED BY ASSOCIATED OVERDOMINANT LOCI IN FINITE POPULATIONS*

TOMOKO OHTA AND MOTOO KIMURA

National Institute of Genetics, Mishima, Japan

Received November 6, 1970

MAINTENANCE of genetic variability is a subject of great interest in population genetics. For more than 15 years, the relative importance of the classical *versus* balance hypotheses (cf. DOBZHANSKY 1955) has been actively discussed among population geneticists.

Recent studies on the isozyme polymorphisms by numerous investigators, following the pioneering works by LEWONTIN and HUBBY (1966) in Drosophila and HARRIS (1966) in man, have increased the acuteness of the problem. This is because the amount of heterozygosity is much higher than previously thought, although this possibility was foreseen and discussed by KIMURA and CROW (1964).

To explain the high level of heterozygosity, several hypotheses have been proposed such as truncation selection with overdominance (SVED, REED and BODMER 1967; KING 1967; MILKMAN 1967), truncation selection with overdominance and linkage (WILLS, CRENSHAW and VITALE 1970), frequency dependent selection (KOJIMA and YARBROUGH 1967), selective neutrality of isoallelic variation (KIMURA 1968a, b), as well as simple overdominance.

Recently KIMURA and OHTA (1971) proposed that random genetic drift of neutral or nearly neutral mutations can account for observed protein polymorphisms. However, some genetic variations, even if a relative minority in terms of the number of loci, must be selected. In this article, we will discuss some consequences of the presence of overdominant loci as influencing the behavior of neutral alleles. The basic idea was derived from our quantitative studies on linkage disequilibrium in finite populations (OHTA and KIMURA 1969a, b, 1970, 1971) which showed that an apparent selective force is created at intrinsically neutral loci by non-random association with selected loci through random drift. Actually, the apparent selection takes the form of heterozygote advantage, i.e., "associative overdominance" as it is named by FRYDENBERG (1963). Therefore sparsely distributed overdominant loci may keep many loci polymorphic, other than themselves. Here, we intend to study the bearing of the associative overdominance on the maintenance of isoallelic variations.

ASSUMPTIONS OF THE MODEL

We assume that intrinsically overdominant loci are sparsely distributed on the whole chromosome, at each of which alleles are segregating around equi-

* Contribution No. 799 from the National Institute of Genetics, Mishima, Shizuoka-ken, 411 Japan. Aided in part by a Grant-in-Aid from the Ministry of Education, Japan.

librium frequencies by sufficiently strong overdominance (heterozygote advantage). We also assume that numerous neutral loci are distributed more densely among them.

Now, our problem is to clarify the behavior of neutral mutants influenced by linkage with overdominant loci. To simplify the treatments we assume that the overdominant loci are equally spaced and are subject to equal selective force with heterozygote advantage s over either homozygote (symmetric overdominance). The following consideration based on the observed inbreeding depression will be helpful to determine the realistic number and selection coefficient of overdominant loci. LATTER and ROBERTSON (1962) studied the decrease in competitive ability in Drosophila at several levels of inbreeding. The three major components measured were: egg laying ability, larval survival and male mating ability. Their data show that the amount of inbreeding depression in competitive ability is expressed approximately as e^{-2F} in which F is the inbreeding coefficient. If we assume that most of the inbreeding depression represents overdominance load, ns amounts to about 4. For example, with 200 loci, this corresponds to 2% overdominance at each locus. If a half of the depression is overdominance load and the remaining half is mutational load, ns becomes 2, and this allows 200 loci with heterozygote advantage of 1%.

In Drosophila the total map length of the genome is about 300 units (in centimorgans). Therefore, on the average, one overdominant locus may exist in approximately one map unit, each with one or two percent overdominance. In addition to overdominant loci, in a much larger number of loci, unconditionally detrimental mutants are segregating at very low frequencies. According to MUKAI (1969) the average number of detrimental mutants per chromosome is only a half dozen. It is possible that most of the inbreeding depression reported by LATTER and ROBERTSON (1962) is due to detrimental mutants. Actually, as long as there exists inbreeding depression, associative overdominance will also be caused by linked detrimentals, but this subject will be treated in a separate paper (OHTA 1971).

It is certain that some portions of genetic variability are selected and that overdominance exists among them. The mutant for sickle cell hemoglobin is a well known example of overdominance. The sickle cell allele has a selective disadvantage by producing anemia in homozygotes but gains an advantage by giving resistance to malaria in heterozygotes. It is likely that this type of ambivalent gene action is the main cause of overdominance (HUXLEY, 1955). When we consider two or more components of fitness, it is possible that a deleterious effect of an allele at each fitness component is recessive so that the overall effect comes out as overdominant. For example, if allele A_1 has selective advantage s_2 over A_2 with almost complete dominance on viability, while A_2 has selective advantage s_1 with almost complete dominance over A_1 on fertility, then the average fitness with both viability and fertility combined will result in overdominance as follows;

genotype	A_1A_1	A_1A_2	A_2A_2
fitness	$1 - s_1$	1	$1 - s_2$

Essentially the same suggestion was made by HALDANE and JAYAKAR (1965) in their discussion on the human genetic load. This type of mechanism works as long as the favorable effect of each allele with respect to an individual fitness component is dominant (not necessarily completely dominant).

It is now known that several kinds of multimer enzymes are formed in the heterozygote and that they sometimes show differential activities (FINCHAM 1966). It is possible that many examples of gene duplication in evolution, such as hemoglobin α, β and γ, owe their establishment to their initial heterozygote advantage (SPOFFORD 1969).

Thus we regard overdominance as a fairly widespread phenomenon and overdominant loci exist sparsely distributed on the whole chromosome. At the same time we assume that a large number of neutral loci is segregating on the same chromosome.

THEORETICAL TREATMENT OF THE MODEL

Before we treat the effect of linked overdominant loci on the behavior of neutral mutants, we will briefly summarize the behavior of neutral isoalleles in a finite population. It is now well known that mutation is caused by changes in DNA base arrangement. With hundreds of nucleotide sites per locus the possible number of allelic states is astronomical. Therefore, in the model of KIMURA and CROW (1964) it is assumed that the number of possible allelic states is so large that whenever a mutant appears it represents a new, not preexisting allele.

Consider a particular locus and let u be the mutation rate per gamete, such that in a population of size N, $2Nu$ neutral mutants are produced each generation. Then the average homozygosity at equilibrium (H_o) is given by

$$H_o = 1/(4N_e u + 1) , \tag{1}$$

where N_e is the effective population number. The reciprocal of this quantity is termed by KIMURA and CROW (1964) the effective number of alleles, n_e, so that $n_e = 4N_e u + 1$. Thus, if $4N_e u$ is larger than 1, the average homozygosity is smaller than 1/2 and the effective number of alleles is larger than 2, while if $4N_e u$ is less than 1, the effective number is less than 2.

If the number of possible allelic states is K (finite) rather than infinite, the expected homozygosity is

$$H_o = \frac{4N_e u\left(\frac{1}{K-1}\right) + 1}{4N_e u\left(\frac{K}{K-1}\right) + 1} . \tag{2}$$

(KIMURA 1968b). Formula (1) is a special case of $K = \infty$. The average heterozygosity ($\bar{H}$) is given by $1 - H_o$. For the special case of $K = 2$, we have

$$\bar{H} = \frac{4N_e u}{8N_e u + 1} , \tag{3}$$

while for $K = \infty$

$$\bar{H} = \frac{4N_e u}{4N_e u + 1} . \tag{4}$$

Since $2Nu$ neutral mutants appear each generation in the population and $2Nu \times 1/2N$ or u of them reach eventual fixation, u represents the rate of mutant substitution (cf. KIMURA 1968a, KING and JUKES 1969, CROW 1969). Thus, the average interval of mutant substitution becomes $1/u$. This means that if $4N_e > 1/u$ or $4N_e u > 1$, the population is segregating most of the time.

Now, our main problem is to estimate the magnitude of associative overdominance at the neutral locus using the model described in the previous section. Let us consider two loci that are separated by recombination fraction c. We will assume that a pair of alleles A_1 and A_2 are maintained by strong overdominance in the first locus, and a steady flux of neutral mutations occur in the second locus. We will denote by p the frequency of A_1 at the first locus and by q the frequency of a mutant at the second, and by D the coefficient of linkage disequilibrium as follows;

frequency $(A_1) = p$	frequency $(A_2) = 1 - p$
frequency $(B_1) = q$	frequency $(B_2) = 1 - q$
frequency $(A_1B_1) = g_1$	frequency $(A_1B_2) = g_2$
frequency $(A_2B_1) = g_3$	frequency $(A_2B_2) = g_4$

$$D = g_1g_4 - g_2g_3$$

To simplify the treatment we assume that overdominance at the first locus is so strong that p is kept at a constant value $\hat{p}$, i.e., $p = \hat{p}$. We will measure the amount of non-random combination of genes between the two loci by

$$\sigma_d^2 = \frac{E\{D^2\}}{E\{p(1-p)q(1-q)\}} \tag{5}$$

which was termed the squared standard linkage deviation by OHTA and KIMURA (1969a). This is analogous to the chi-square contingency coefficient used in statistics. As shown by OHTA and KIMURA (1971b), under steady flux of mutations we have

$$E\{L(f)\} + \Delta_{mut}E(f) = 0 , \tag{6}$$

where L is the differential operator of the Kolmogorov backward equation involved, E is the expectation operator, f is a polynomial in p, q and D, and vanishes at the periphery of the square $0 \leqq p \leqq 1$, $0 \leqq q \leqq 1$, and $\Delta_{mut}E(f)$ is the constant input per generation by new mutations with respect to $E(f)$.

For the present model, $p\ (= \hat{p})$ is constant and the differential operator is given by

$$L = \frac{1}{4N_e}\left[q(1-q) - \frac{D^2}{\hat{p}(1-\hat{p})}\right]\frac{\partial^2}{\partial q^2} + \frac{1}{2N_e}\left[(1-2q)D + \frac{2\hat{p}-1}{\hat{p}(1-\hat{p})}D^2\right]\frac{\partial^2}{\partial q \partial D}$$

$$+\frac{1}{4N_e}\left[\hat{p}(1-\hat{p})q(1-q)+(1-2\hat{p})(1-2q)D\right.$$
$$\left.-\frac{1-3\hat{p}(1-\hat{p})}{\hat{p}(1-\hat{p})}D^2\right]\frac{\partial^2}{\partial D^2}-cD\frac{\partial}{\partial D} \tag{7}$$

This operator is equivalent to the differential operator of equation (9) of OHTA and KIMURA (1970) except for the factor $1/N_e$ and except that this operator does not contain terms giving the effect of mutation. In the present case, the effect of mutation is given by $\Delta_{\text{mut}}E(f)$. If we let $f=2q(1-q)$ in equation (6) we get

$$2E[q(1-q)]=\frac{2N_e v_m}{N}+\frac{2E(D^2)}{\hat{p}(1-\hat{p})}, \tag{8}$$

where v_m is the number of new mutants that appear at the second locus in each generation such that $v_m=2Nu$ and we take $\Delta_{\text{mut}}E(f)=2v_m/(2N)$. Notice that $2E[q(1-q)]$ gives the average number of heterozygous nucleotide sites at the second locus per individual. Next, let $f=qD$, then,

$$\frac{2\hat{p}-1}{\hat{p}(1-\hat{p})}E(D^2)=2(N_ec+1)E(qD)\ . \tag{9}$$

Finally, let $f=D^2$, we get

$$\hat{p}(1-\hat{p})\,E\{q(1-q)\}-2(1-2\hat{p})E(qD)$$
$$-\left[4N_ec+\frac{1-3\hat{p}(1-\hat{p})}{\hat{p}(1-\hat{p})}\right]E(D^2)=-\frac{N_ev_m}{2N^2} \tag{10}$$

By solving the three simultaneous equations (8), (9) and (10), we get,

$$\sigma_d^2=\frac{E(D^2)}{\hat{p}(1-\hat{p})\,E[q(1-q)]}=\frac{1+\dfrac{1}{2N}}{4N_ec+\dfrac{4}{N_ec+1}-3+\dfrac{N_ec}{\hat{p}(1-\hat{p})\,(N_ec+1)}+\dfrac{1}{2N}}$$
$$\approx\frac{1}{4N_ec+\dfrac{4}{N_ec+1}-3+\dfrac{N_ec}{\hat{p}(1-\hat{p})\,(N_ec+1)}}\ . \tag{11}$$

If N_ec is much larger than unity, we have

$$\sigma_d^2\approx\frac{1}{4N_ec} \tag{12}$$

with good approximation. It is remarkable that this approximation for a large N_ec is valid for all cases so far investigated (assuming no epistasis), that is, the state of steady decay, stationary state with recurrent mutation and/or overdominance, and also the state of steady flux of neutral mutations (OHTA and KIMURA 1969a, b, 1970, 1971), as well as the present case.

In terms of σ_d^2, the apparent heterozygote advantage at the neutral locus is approximately (OHTA and KIMURA, 1970a),

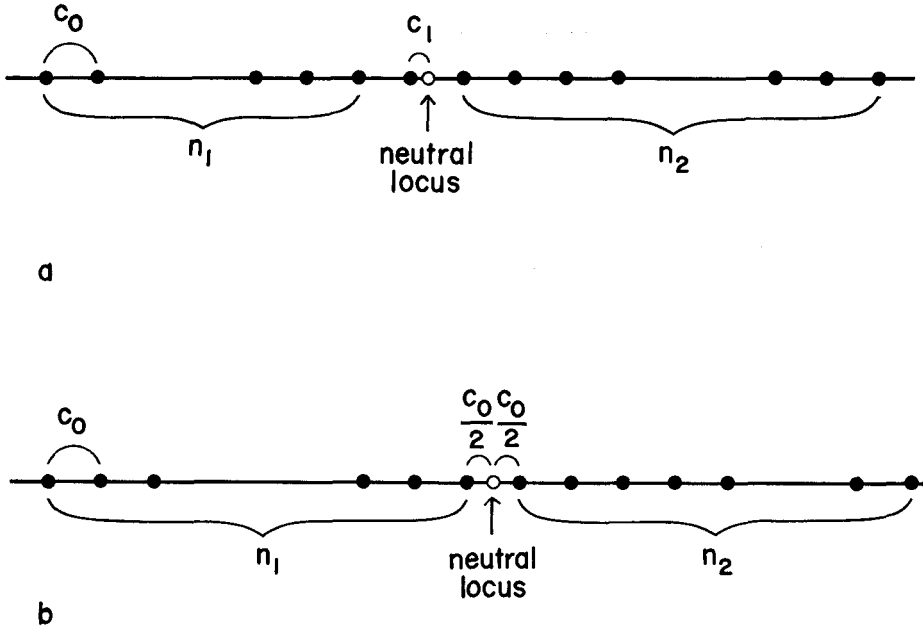

FIGURE 1.—Two models used to investigate associative overdominance. In 1a, the neutral locus is located near one of the overdominant loci (case 1), while in 1b the neutral locus is located equal distance between adjacent overdominant loci (case 2).

$$E\{w_{B_1B_2} - w_{B_1B_1}\} \approx (s_1 + s_2)\sigma_d^2\left[\frac{\hat{p}(1-\hat{p})}{q}\right]$$
$$E\{w_{B_1B_2} - w_{B_2B_2}\} \approx (s_1 + s_2)\sigma_d^2\left[\frac{\hat{p}(1-\hat{p})}{1-q}\right] . \quad (13)$$

In these formulae, s_1 and s_2 stand for selection coefficients against two homozygotes, A_1A_1 and A_2A_2, and we assume that the frequency of mutants at the neutral locus happens to take an intermediate frequency q not very far from 0.5.

In order to extend the above results to multi-locus overdominance, we assume that the epistasis coming from multiplicative fitness effect among loci is not strong enough in comparison with the recombination fraction between adjacent overdominant loci so that a super gene of the type discussed by FRANKLIN and LEWONTIN (1970) is not formed. In the following treatments, we take logarithms of individual fitness values for a multiplicative overdominance. Then the selection coefficient of the associative overdominance which we denote by s' becomes in general the sum of the effects of all linked overdominant loci. Assuming $s_1 = s_2 = s$ (symmetric overdominance) for all overdominant loci and also assuming $q \approx 0.5$ for a particular neutral loci, we will consider two simple cases. First, if the neutral locus happens to be very near one of the overdominant loci (case 1) as illustrated in Figure 1a, the coefficient of associative overdominance becomes

$$s' = \frac{s}{4N_e c_1} + \frac{s}{4N_e c_0} \{\log_e n_1 + \log_e n_2 + 2\gamma\} , \qquad (14)$$

where γ is Euler's constant (0.577), c_0 is the recombination fraction between two adjacent overdominant loci and c_1 is that between the neutral and the nearest overdominant locus. In deriving this equation we assume that $c_1 << c_0$ and $4N_e c_1 >> 1$. The first term on the right-hand side of the equation represents the effect of the nearest overdominant locus and it is larger than the effect of the remaining $(n_1 + n_2)$ loci combined since c_1 is much smaller than c_0. Secondly, if the neutral locus is located at an equal distance between adjacent overdominant loci (case 2) as illustrated in Figure 1b, s' becomes,

$$s' = \frac{s}{4N_e c_0} \{\log_e n_1 + \log_e n_2 + 2\gamma + 4\log_e 2\} . \qquad (15)$$

This is smaller than the corresponding value given by (14) because no overdominant locus is tightly linked to the neutral locus.

Let us consider some examples. Suppose there are 100 overdominant loci in a chromosome one morgan in length. If the neutral locus is located at about the middle of the chromosome very near to the 50th overdominant locus with $c_1 = 10^{-3}$, then taking $c_0 = 0.01$, $n_1 = 49$, $n_2 = 50$ in formula (14), we obtain $N_e s' \approx s/2 \times 10^3$ for a neutral mutant allele with frequency near 1/2. On the other hand, if a neutral locus is located just in the middle between the 50th and 51st overdominant loci (case 2), then $N_e s'$ becomes approximately $s/4 \times 10^3$ from formula (15). Thus, the value of $N_e s'$ heavily depends on the nearest overdominant locus. However, as pointed out by OHTA and KIMURA (1970) and KIMURA (1971), the value of $N_e s'$ is independent of the effective population size, N_e. Also it is not much influenced by the number of overdominant loci as long as the total overdominant load per unit length of chromosome is constant because s/c_0 remains constant.

From these examples, we note that associative overdominance plays an important role only when the population size is of the order of 10^3 or less. However, such a small population size is realistic since the effective population size applicable in the present treatment is the local size rather than that of the whole species. This is because the migration rate is usually much smaller than the recombination fraction and the effect of crossing over will predominate that of migration. A more detailed discussion of this subject will be given elsewhere.

In the above discussion, we assumed that the frequency at the neutral or the selected loci takes an intermediate value. If it takes an extreme value either near 0 or 1, a different treatment is necessary. The analysis of the associative overdominance due to linked detrimentals will be given in OHTA (1971).

Finally let us consider briefly the influence of the associated overdominant loci on the heterozygosity at the neutral locus. We note that equation (8) gives the average number of heterozygous nucleotide sites at the neutral locus. Then the first term on the right hand side gives the number expected under neutral mutation while the second term gives the amount of increase due to association with

an overdominant locus. Therefore, if we denote the number of heterozygous nucleotide sites in the second locus by $\bar{H}'$ then,

$$\bar{H}' = 2E[q(1-q)] = \frac{2N_e v_m}{N} + 2E\{q(1-q)\}\sigma_d^2$$

$$\approx \bar{H}'_n [1 + \sigma_d^2] \tag{16}$$

where $\bar{H}'_n$ is the value of $\bar{H}'$ expected without the effect of associated overdominant loci. In other words, heterozygosity will be enhanced by the fraction σ_d^2 by associative overdominance. If there are many overdominant loci, their effect on enhancing heterozygosity would be proportional to the sum of σ_d^2 as long as it is small. At present, however, we are not quite sure of this effect on enhancing heterozygosity. The reason is that when a mutant frequency is small, the large variance of the apparent selective value at the neutral locus will reduce the heterozygosity. Associative overdominance clearly retards the fixation of neutral mutants having intermediate frequencies, although we are not sure of its effect on increasing the overall heterozygosity.

MONTE CARLO EXPERIMENTS

In order to verify the results of the theoretical treatments given in the previous section, extensive Monte Carlo experiments were performed simulating a multilocus Genetic system, using a TOSBAC 3400 computer at the National Institute of genetics. The simulated population consists of N diploid individuals each having two homologous chromosomes represented by binary integers. In this model there are only 2 alleles (represented by numerals *0* or *1*) at each locus.

Each generation consists of mutation, crossing over, selection and sampling. Mutation is simulated by changing 0 to 1 or vice versa at a locus and the decision of occurrence and the determinations of an individual chromosome and locus were made by generating pseudo-random numbers using subroutine RAND 20 in FORTRAN IV. The mutation rates are equal in both directions, since if a gene randomly chosen happens to be *0*, it mutates to *1* and vice versa.

The decision of having crossing over and the determination of the position of exchange were again made by generating pseudo-random numbers. We assumed equal crossing over between adjacent loci. Sampling with selection was carried out as follows: One individual is randomly chosen and its number of homozygous loci m is determined. Then a uniform random number between 0 and 1 was produced and if it happens to be smaller than e^{-sm}, one of two chromosomes of this individual was randomly sampled, otherwise we proceed to the next randomly chosen individual. This process is repeated until $2N$ chromosomes are sampled to form the next generation. This model of selection corresponds to multiplicative overdominance between loci.

Three different types of experiments were performed. In the first experiment, the coefficient of associative overdominance was studied. Each chromosome has 23 loci and the recombination fraction between two adjacent loci c_0 ranges from

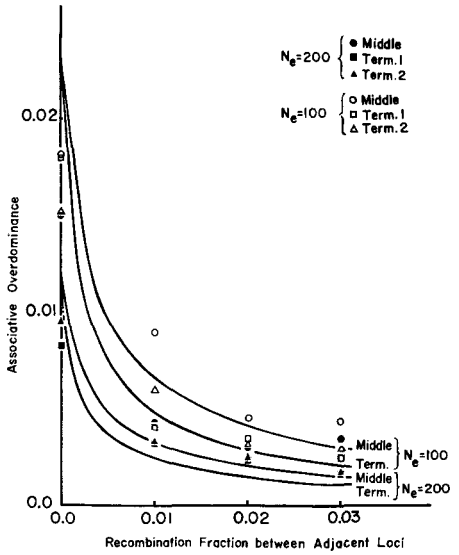

FIGURE 2.—Associative overdominance observed at the three marker loci in Monte Carlo experiments. The coefficient of associative overdominance is shown as a function of recombination fraction. Ordinate: the coefficient of associative overdominance s'. Abscissa: recombination fraction between adjacent loci (c_0). Curves represent theoretical predictions (formulae 17 or 18), while dots give observed values. Parameters in the experiments are; $s = 0.01$, $u = 0.005$, $N_e = 100$ or 200, and the total number of loci is 23. Each experimental value is the average over 500 generations.

0.0 to 0.03. The mutation rate is 0.005 per locus per generation. Two terminal loci, "terminal 1" and "terminal 2", and the middle locus are used as markers, of which terminal 1 and the middle loci are neutral. All the remaining 21 loci are overdominant with $s_1 = s_2 = 0.01$ and the population size is either 100 or 200. For this type of model, the selection coefficient of associative overdominance at a selected locus predicted from the theory is $s + s'$ in which s' is given by

$$s' = \frac{s}{4N_e c_0} \{\log_e n_1 + \log_e n_2 + 2\gamma\} , \tag{17}$$

where n_1 and n_2 are the numbers of overdominant loci on the left and the right of that locus (OHTA and KIMURA 1970). For a terminal locus, one of n_1 and n_2 is zero, and we have, assuming $n_2 = 0$,

$$s' = \frac{s}{4N_e c_0} (\log_e n_1 + \gamma) . \tag{18}$$

Figure 2 shows the observed values (dots) and the theoretical predictions (curves) of associative overdominance (s') at the three marker loci. Although the theoretical prediction seems to underestimate the true value slightly, the agreement between the two is satisfactory. However, with much stronger selection and/or tighter linkage, the effect reported by FRANKLIN and LEWONTIN

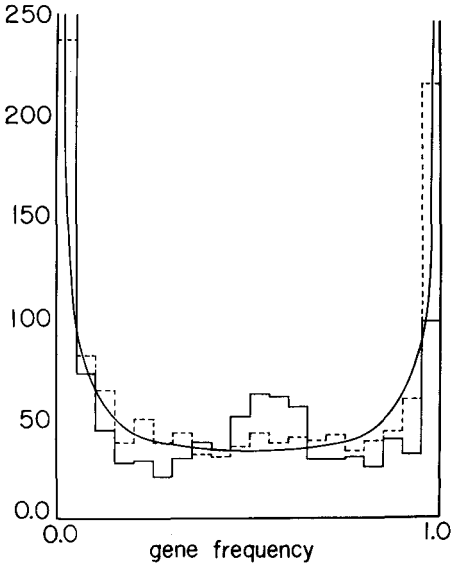

FIGURE 3.—Distribution of gene frequencies at the neutral marker loci (the middle and "terminal 1" loci). Histograms with solid lines show the observed frequency distribution at the middle locus and those with dotted lines at the terminal locus. The curve represents the theoretical prediction under recurrent mutation with rate $N_e u = 0.1$, disregarding the effects from other loci. The experimental distributions are the results of simulation experiments extending over 12,500 generations with output at every 10 generations. Parameters are; $s = 0.05$, $c_0 = 0.01$, $N_e = 100$ and $u = 0.001$ (two marker loci) or $u = 0.005$ (remaining loci).

(1970) (formation of a supergene by epistasis) appears and our formulae become less satisfactory.

In the second experiment, the gene frequency distributions at the middle and the terminal neutral loci were investigated using the same model as before. In this case, the selection coefficient at the overdominant loci was 5% instead of 1%, $N_e = 100$, $c_0 = 0.01$ and u (mutation rate per locus) were the same as before. However, the mutation rates at the neutral loci are lowered to $N_e u = 0.1$ to make the detection of associative overdominance easier. The results are shown in Figure 3. The curve represents the theoretical distribution at the neutral locus under recurrent mutation with rate $N_e u = 0.1$, disregarding the effects from other loci. The observed frequency distribution is shown by histograms with dotted lines for the terminal locus (terminal 1) and solid lines for the middle locus. They represent the results of experiments over 12,500 generations with outputs at every 10th generation.

As may be seen from Figure 3, the associative overdominance is effective in increasing the chance of a neutral mutant staying at a gene frequency of around 50% at the middle locus but not at the terminal locus. Theoretically, it is expected that $N_e s'$ is about 3.5 at the middle locus and about 2.5 at the terminal locus. $N_e s'$ of 3.5 seems to be equivalent in effect to the true overdominance of $N_e s \approx 1$, and $N_e s'$ of 2.5 appears not to be effective at all. The average hetero-

TABLE 1

Observed and expected heterozygosity in the second experiments

Expected heterozygosity	Observed heterozygosity	
From equation (3)	Middle locus	Terminal locus
0.222	0.217	0.226

zygosity over 12,500 generations at these loci was also checked in this experiment and it turned out to be just as expected from the neutral mutation theory as shown in Table 1.

In the third experiment, the effect of associative overdominance on retarding fixation was measured using the same model and the same parameters as the second experiment. In this experiment, however, the program was modified so that whenever loss or fixation occurs at each of the neutral marker loci, the mutant frequency at that locus is set equal to 0.5. No recurrent mutations are assumed at these loci. Thus, the numbers of generations until fixation or loss starting from initial frequency 0.5 were recorded. The number of generations until fixation or loss was 398 ± 70 (average of 27 replications) at the middle locus and 325 ± 45 (average of 33 replications) at the terminal ("terminal 1") locus. On the other hand, the theoretical value using the formula by KIMURA and OHTA (1969) for $p = 0.5$ is

$$\bar{t}_1(p) = \bar{t}_0(p) = 4N_e \log_e 2 = 277 ,$$

since $N_e = 100$. Thus, again the associative overdominance is effective only at the middle locus and it corresponds to true overdominance of only $N_e s \approx 1$ as before (cf. ROBERTSON 1962, KIMURA and OHTA 1969).

We have checked only a limited number of cases. However, it is clear that associative overdominance exists under our model and it retards the fixation of the neutral mutants having intermediate frequencies. From the experiments it appears as if the associative overdominance of a given $N_e s'$ has effect only some one fourth as large as the true overdominance having the same value of $N_e s$, but $N_e s'$ may take a larger value in actual populations than in the present simulated populations.

DISCUSSION

It is possible that the pattern of observed protein polymorphisms can be explained, to the first approximation, by the neutral mutation-random drift theory (cf. KIMURA and OHTA 1971). However, not all loci are selectively neutral, and as long as there exists inbreeding depression, we have to take into account the effect of selected loci.

We believe that the associative overdominance discussed in this article will influence the behavior of the neutral mutant alleles. Although we are not sure if it enhances the average (overall) heterozygosity, it certainly retards the fixation of mutants with intermediate gene frequencies. It is not as effective as true

overdominance, but $N_e s'$ could take a large value. If we use the example discussed in the section on the model, $N_e s'$ amounts to $s/2 \times 10^3 \sim s/4 \times 10^3$. Thus with 2% overdominance at each locus ($s = 0.02$), which is acceptable from the standpoint of load, $N_e s'$ becomes $10 \sim 5$. In the case of Drosophila, because of no crossing over in males, $N_e s'$ becomes $20 \sim 10$ with the same parameters. The important point is that this value is not much influenced by the number of overdominant loci as long as the overdominance load per unit chromosome is constant.

Such associative overdominance must be effective in retarding fixation of segregating alleles in semi-isolated local populations. When a small group of individuals is semi-isolated from the main body of the species, the variation initially present would persist longer due to associative overdominance than without it. Together with the migration effect, this will bring about the constant distribution of allelic frequencies throughout the entire species. The associative overdominance is probably ineffective in increasing overall heterozygosity of the species, although it may influence the frequency distribution of the neutral alleles as shown in Figure 3.

We have emphasized in the section regarding the model that there are many components of fitness and overdominance in overall fitness may be produced at some loci, even if there is only slight dominance for each fitness component. HALDANE and JAYAKER (1965) pointed out that differential gene action on mortality, fertility or disease resistance is the cause of overdominance in many cases. Even male and female fertilities are usually different components of fitness. We consider that genes with small effects may be important in this respect, since genes with drastic effects on viability or fertility are mostly balanced by mutation and selection. In this respect, the sickle cell hemoglobin seems to represent an exceptional case.

Experimental measures of fitness such as relative viability and competitive ability represent mostly particular components or a combination of a few components of the real fitness. With such measures, it may often be difficult to estimate the actual overdominance load. Inversion polymorphisms in Drosophila show clear overdominance on viability. This must be the sum of the effects of many loci within the inversion, each having some dominance on viability.

If the overdominance load is the same in Drosophila and in man, associative overdominance is more pronounced in Drosophila than in man, since the Drosophila genome has fewer chromosomes and also no recombination occurs in males. Furthermore, inversions might be more common in Drosophila than in man.

Thus, we may infer that most of the isozyme polymorphisms are intrinsically neutral or nearly so but a relatively small fraction are overdominant. We do not know the actual number of overdominant and neutral loci. There might be many overdominant loci with very small effects. The important point is that with a tolerable amount of overdominance load, the observed genetic variations can easily be explained. We must add also that when a small number of chromosomes are extracted from natural populations and rapidly multiplied for an experiment, stronger associative overdominance will result.

Another important point is that associative overdominance does not change the long term average of the rate of substitution of neutral mutants by random drift. It may increase the variance of the evolutionary rates among lines, sometimes favoring and sometimes disfavoring individual mutants to multiply.

SUMMARY

The behavior of isoallelic mutants in finite populations was investigated with special reference to nonrandom association of neutral isoallelic mutants to overdominant loci by random genetic drift. The overdominant loci are assumed to be sparsely distributed along the whole chromosomes and the neutral loci are distributed more densely among them. The behavior of neutral mutants is influenced by the surrounding overdominant loci and the apparent selective force takes the form of "associative overdominance". This was treated theoretically using a model assuming that the overdominant loci are equally spaced and have equal selection coefficients. For this model the approximate magnitude of associative overdominance was estimated. Monte Carlo experiments proved the validity of the theoretical prediction, although the estimated degree of associative overdominance is not as effective as true overdominance. For the set of parameters used in the experiments, it was about 1/4 as effective as true overdominance in retarding gene fixation. The associative overdominance at intrinsically neutral loci, will contribute, at least partly, to bring about a constant distribution of neutral alleles, by preventing fixation of these alleles in local populations.

LITERATURE CITED

CROW, J. F., 1969 Molecular genetics and population genetics. Proc. 12th Intern. Congr. Genet. **3**: 105–113.

DOBZHANSKY, TH., 1955 A review of some fundamental concepts and problems of population genetics. Cold Spring Harbor Symp. Quant. Biol. **20**: 1–15.

FINCHAM, J. R. S., 1966 *Genetic Complementation.* Benjamin, New York.

FRANKLIN, I. and R. C. LEWONTIN, 1970 Is the gene the unit of selection? Genetics **65**: 707–734.

FRYDENBERG, O., 1963 Population studies of a lethal mutant in *Drosophila melanogaster* I. Behaviour in populations with discrete generations. Hereditas **50**: 89–116.

HALDANE, J. B. S. and S. D. JAYAKAR, 1965 The nature of human genetic loads. J. Genet. **59**: 143–149.

HARRIS, H., 1966 Enzyme polymorphism in man. Proc. Roy. Soc. London B; **164**: 298–310.

HUXLEY, J., 1955 Morphism and evolution. Heredity **9**: 1–52.

KIMURA, M., 1968a Evolutionary rate at the molecular level. Nature **217**: 624–626. ——, 1968b Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. Genet. Res. **11**: 247–269. ——, 1971 Theoretical foundation of population genetics at the molecular level. Theoretical Population Biology **2**: 174–208.

KIMURA, M. and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49**: 725–738.

KIMURA, M. and T. OHTA, 1969 The average number of generations until fixation of a mutant

gene in a finite population. Genetics **61**: 763–771. ——, 1971 Protein polymorphism as a phase of molecular evolution. Nature **229**: 467–469.

KING, J. L., 1967 Continuously distributed factors affecting fitness. Genetics **55**: 483–492.

KING, J. L. and T. H. JUKES, 1969 Non-Darwinian evolution: Random fixation of selectively neutral mutations. Science **164**: 788–798.

KOJIMA, K. and K. M. YARBROUGH, 1967 Frequency-dependent selection at the esterase-6 locus in *Drosophila melanogaster*. Proc. Nat. Acad. Sci. U.S. **57**: 645–649.

LATTER, B. D. H. and A. ROBERTSON, 1962 The effects of inbreeding and artificial selection on reproductive fitness. Genet. Res. **3**: 110–138.

LEWONTIN, R. C. and J. L. HUBBY, 1966 A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics **54**: 595–609.

MILKMAN, R. D., 1967 Heterosis as a major cause of heterozygosity in nature. Genetics **55**: 493–495.

MUKAI, T., 1969 Maintenance of polygenic and isoallelic variation in populations. Proc. XII Intern. Congr. Genetics Vol. **3**: 293–308.

OHTA, T., 1971 Associative overdominance caused by linked detrimental mutations. Genet. Res. (in press).

OHTA, T. and M. KIMURA, 1969a Linkage disequilibrium due to random genetic drift. Genet. Res. **13**: 47–55. ——, 1969b Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics **63**: 229–238. ——, 1970 Development of associative overdominance through linkage disequilibrium in finite populations. Genet. Res. **16**: 165–177. ——, 1971 Linkage disequilibrium between two segregating nucleotide sites under steady flux of mutations in a finite population. Genetics, (in press).

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

SPOFFORD, J. B., 1969 Heterosis and the evolution of duplications. Am. Naturalist **103**: 407–432.

SVED, J. A., T. E. REED and W. F. BODMER, 1967 The number of balanced polymorphisms that can be maintained in a natural population. Genetics **55**: 469–481.

WILLS, C., J. CRENSHAW and J. VITALE, 1970 A computer model allowing maintenance of large amounts of genetic variability in Mendelian populations. I. Assumptions and results for large populations. Genetics **64**: 107–123.