MODIFICATION OF LINKAGE INTENSITY BY NATURAL SELECTION

MASATOSHI NEI

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

Received June 8, 1967

GENETIC linkage or linkage of genes is one of the common genetic features of all organisms from viruses to man. It may be of great concern to evolutionary geneticists how genetic linkage was developed and how the intensity of linkage has been adjusted in the evolutionary process. FISHER (1930) suggested that "the presence of pairs of factors in the same chromosome, the selective advantage of each of which reverses that of the other, will always tend to diminish recombination, and therefore, to increase the intensity of linkage in the chromosomes of that species." He also recognized another agency which would increase recombination, namely, the constant spread of advantageous mutations which, unless they occur so seldom that each has become predominant before the next appears, can only come together in the same gamete by means of recombination. This problem was recently discussed at length but still semi-quantitatively by BODMER and PARSONS (1962).

There is a large amount of evidence that linkage intensity is under genetic control (cf. BODMER and PARSONS 1962). In *Drosophila ananassae*, for example, MORIWAKI (1940) identified a dominant gene, *En-2*, located on the right arm of the second chromosome, which enhances recombination between almost every pair of loci on the same chromosome in both males and females, while KIKKAWA (1937) found another gene (or genes) which induces crossing over in males in the third chromosome. CLARK and MARGULIES (1965) and HOWARD-FLANDERS and THERIOT (1966) also reported several recombination-deficient mutants in *Escherichia coli* K-12, which are probably due to single-gene mutation. Further, the effectiveness of artificial selection in reducing or increasing recombination values was reported by DETELFSEN and ROBERTS (1921), PARSONS (1958), and MUKHERJEE (1961) in Drosophila.

The purpose of this paper is to present mathematical models for the development of genetic linkage and the modification of linkage intensity by natural selection. In the present paper we shall consider only those agencies which would increase the linkage intensity. KIMURA (1956) developed a model of a genetic system which leads to closer linkage by natural selection in diploid organisms, but he was concerned mainly with a crossover-reducing mechanism such as an inversion.

[1] This investigation was supported in part by a grant from the Scientific Research Fund of the Ministry of Education, Japan, and also by a grant from the Toyo Rayon Foundation.

Haploid system

Model 1: The modification of linkage intensity may occur in two different manners, namely, through structural changes of chromosomes and through selection of modifier genes. When a structural change such as an inversion or translocation occurs, the genes which were located far apart on the same chromosome or on different chromosomes before the change, may come close together. The present model represents this type of modification. Structural changes of chromosomes are known to occur in fungi (cf. BARRY 1967) and in bacteria (JACOB and WOLLMAN 1961; BECKWITH and SIGNER 1966). It should also be noted that a certain type of development of new linkage may be represented by this model, as will be discussed later.

Consider two pairs of genes each with two alleles, *A-a* and *B-b*, and suppose that these two pairs of genes are linked with a recombination value r, and that the fitnesses and the frequencies before selection of the four possible genotypes are as follows:

Genotype	AB	Ab	aB	ab
Fitness	W_1	W_2	W_3	W_4
Frequency	p_1	p_2	p_3	p_4

As shown by FELSENSTEIN (1965) and KIMURA (1965), the amounts of changes in p_1, p_2, p_3, and p_4 per generation are then given by

$$\begin{aligned}\Delta p_1 &= p_1(W_1 - \overline{W})/\overline{W} - rD\\ \Delta p_2 &= p_2(W_2 - \overline{W})/\overline{W} + rD\\ \Delta p_3 &= p_3(W_3 - \overline{W})/\overline{W} + rD\\ \Delta p_4 &= p_4(W_4 - \overline{W})/\overline{W} - rD\end{aligned}$$

respectively, where $\overline{W} = \sum_{i=1}^{4} p_i W_i$ and $D = (p_1 W_1 p_4 W_4 - p_2 W_2 p_3 W_3)/\overline{W}^2$.

Now suppose that a new completely linked gene pair, $\hat{ab}$, arises from ab by some mechanism, and let x and $p_4{}^u$ be the frequencies of $\hat{ab}$ and ab, respectively, with $x + p_4{}^u = p_4$. We assume that the fitness of $\hat{ab}$ remains the same as that of ab. Then we have

$$\begin{aligned}\Delta p_4{}^u &= p_4{}^u(W_4 - \overline{W})/\overline{W} - rD^u\\ \Delta x &= x(W_4 - \overline{W})/\overline{W}\end{aligned} \qquad (1)$$

where $D^u = (p_1 W_1 p_4{}^u W_4 - p_2 W_2 p_3 W_3)/\overline{W}^2$. Hence, genotype $\hat{ab}$ has a selective advantage over ab, if rD^u is positive. This quantity or D^u is expected to be positive only when the coupling genotypes are fitter than the repulsion ones. When D^u or D is negative, those linkages which occur in Ab or aB have a selective advantage. If there is no epistasis, D or D^u is expected to be 0 at least after several cycles of random mating, so that $\hat{ab}$ has no selective advantage over ab.

Anyway, if $W_4 - \overline{W}$ is positive, genotype $\hat{ab}$ is expected to increase in frequency. In the initial stage of increase in x, $\overline{W}$ is almost constant if the differences between W's are small. In this situation,

$$\Delta x/x = (W_4 - \hat{\overline{W}})/\hat{\overline{W}} \equiv \alpha \qquad (2)$$

where $\hat{\overline{W}}$ is the equilibrium value of $\overline{W}$ before the introduction of $\hat{ab}$. Therefore,

the frequency of $\hat{ab}$ in the initial stage of its increase is approximately given by

$$x_n = x_0 e^{an} \tag{3}$$

where x_n denotes the value of x in the nth generation.

As an example, consider the case where $r = 0.5$, $W_1 = 1$, $W_2 = 1 - s_b$, $W_3 = 1 - s_a$, and $W_4 = 1 + t$, s_a, s_b, and t being small positive constants. In the absence of linked gene pair $\hat{ab}$, there is an unstable equilibrium for the gene frequency of a or b, below which these gene frequencies or p_4 cannot increase even if ab is the fittest genotype. The equilibrium values of gene frequencies of a and b are approximately given by $s_b/(s_a + s_b + t)$ and $s_a/(s_a + s_b + t)$ respectively (CROW and KIMURA 1965). Below these equilibrium values, the genes a and b are maintained only by mutation, so that the frequency of ab is very small. However, once the linked gene pair $\hat{ab}$ is introduced, it is expected to increase rapidly after some period of stochastic changes. In this case

$$\Delta x = x(1-x)t/\overline{W}$$
$$\approx x(1-x)t$$

Hence, the frequency of $\hat{ab}$ in the nth generation is approximately given by

$$x_n = x_0 e^{tn}/(1 - x_0 + x_0 e^{tn}) \tag{4}$$

Therefore, x_n increases logistically. For example, if $t = 0.06$ and $x_0 = 0.001$, about 230 generations are required for x_n to reach 0.99.

Model 2: In this model it is assumed that the recombination value between a pair of loci is controlled by another modifier locus and the modification of linkage intensity is brought about by an increase in the frequency of the modifier gene. Consider two loci each with two alleles, *A-a* and *B-b*, and suppose that the recombination value between these two loci is modified by another locus, i.e. *M-m*. For simplicity we assume that this modifier locus is independent of the other two loci and has no biological function other than modifying the recombination value. For the loci A and B there are four possible genotypes, i.e., AB, Ab, aB, and ab, the fitnesses of which are denoted again by W_1, W_2, W_3 and W_4, respectively. Let the frequencies of AB, Ab, aB, and ab *after selection* in a generation be P_1, P_2, P_3, and P_4 respectively ($P_1 + P_2 + P_3 + P_4 = 1$), and the frequencies of genes M and m be x and y ($= 1 - x$). Note that $P_i = p_i W_i/\overline{W}$. Then, under random mating the frequencies of genotypes ABM, ABm, AbM, etc. will be P_1x, P_1y, P_2x, etc., respectively.

There are ten possible mating types for the loci A and B, as given in Table 1. Each of these mating types can be subdivided into three groups according to the mating type for the modifier locus, i.e. $M \times M$, $M \times m$, and $m \times m$ with frequencies x^2, $2xy$, and y^2 respectively. Except for the two mating types $AB \times ab$ and $Ab \times aB$, this subdivision is irrelevant to the production of genotypes in the next generation. With $AB \times ab$ and $Ab \times aB$ the frequencies of offspring genotypes are dependent on the recombination value, and let the recombination values for matings $M \times M$, $M \times m$, and $m \times m$ be r_2, r_1, and r_0, respectively, with $r_2 \leqslant r_1 \leqslant r_0$.

In bacteria, the recombination of genes is accomplished by any of several processes—sexual mating, transduction, and transformation, but the recombination at a particular time is almost always restricted to a short segment of chro-

TABLE 1

Frequencies of mating types and genotypes produced in model 2

Mating type	Frequency	Genotype produced			
		AB	*Ab*	*aB*	*ab*
$AB \times AB$	P_1^2	1	..	..	..
$AB \times Ab$	$2P_1P_2$	½	½	..	..
$AB \times aB$	$2P_1P_3$	½	..	½	..
$AB \times ab$	$2P_1P_4$				
$M \times M$	x^2	$(1-r_2)/2$	$r_2/2$	$r_2/2$	$(1-r_2)/2$
$M \times m$	$2xy$	$(1-r_1)/2$	$r_1/2$	$r_1/2$	$(1-r_1)/2$
$m \times m$	y^2	$(1-r_0)/2$	$r_0/2$	$r_0/2$	$(1-r_0)/2$
$Ab \times Ab$	P_2^2	1	..	..	..
$Ab \times aB$	$2P_2P_3$				
$M \times M$	x^2	$r_2/2$	$(1-r_2)/2$	$(1-r_2)/2$	$r_2/2$
$M \times m$	$2xy$	$r_1/2$	$(1-r_1)/2$	$(1-r_1)/2$	$r_1/2$
$m \times m$	y^2	$r_0/2$	$(1-r_0)/2$	$(1-r_0)/2$	$r_0/2$
$Ab \times ab$	$2P_2P_4$	..	½	..	½
$aB \times aB$	P_3^2	..	..	1	..
$aB \times ab$	$2P_3P_4$	..	..	½	½
$ab \times ab$	P_4^2	..	..	..	1

mosome. If, therefore, the modifier locus is located far-off from the A and B loci, the pairing of modifier genes does not usually take place in association with the pairing of the A and B loci. Thus, each mating type for the A and B loci is subdivided into only two groups, M and m, which show recombination values of, say, r'_1 and r'_0 respectively. However, this situation can be covered by the above treatment with a simple transformation $r'_0 - r'_1 = 2(r_2 - r_1) = 2(r_1 - r_0)$.

Since the modifier locus is independent of the other two loci, the frequency of genotype ABM before selection in the next generation becomes
$x[P_1^2 + P_1P_2 + P_1P_3 + P_1P_4(1 - xr_2 - yr_1) + P_2P_3(xr_2 + yr_1)] = x(P_1 - r_MD)$
where $r_M = xr_2 + yr_1$ and $D = P_1P_4 - P_2P_3$, which is identical to the previous D. For other genotypes we have

Genotype	Frequency
AbM	$x(P_2 + r_MD)$
aBM	$x(P_3 + r_MD)$
abM	$x(P_4 - r_MD)$
ABm	$y(P_1 - r_mD)$
Abm	$y(P_2 + r_mD)$
aBm	$y(P_3 + r_mD)$
abm	$y(P_4 - r_mD)$

where $r_m = xr_1 + yr_0$. Thus, the average fitness of genotype M becomes

$$\begin{aligned}\overline{W}_M &= (P_1 - r_MD)W_1 + (P_2 + r_MD)W_2 + (P_3 + r_MD)W_3 + (P_4 - r_MD)W_4 \\ &= W_E - r_MD\varepsilon \end{aligned} \quad (5)$$

where W_E is $\sum_{i=1}^{4} P_iW_i$ and $\varepsilon = W_1 - W_2 - W_3 + W_4$. On the other hand, the fitness of genotype m is given by

$$\overline{W}_m = W_E - r_mD\varepsilon \quad (6)$$

and the average fitness of population by

$$\begin{aligned}\overline{W} &= x\overline{W}_M + y\overline{W}_m \\ &= W_E - \bar{r}D\varepsilon \end{aligned} \tag{7}$$

where $\bar{r} = xr_M + yr_m$.

Therefore, the genotype or gene frequency of M after selection is given by

$$x' = x\overline{W}_M/\overline{W} \tag{8}$$

and the amount of change in gene frequency per generation becomes

$$\Delta x = (xy/\overline{W})(r_m - r_M)D\varepsilon\,. \tag{9}$$

Thus, if D and ε are both positive or both negative, the frequency of the modifier gene always increases.

In an arbitrary random mating population, the signs of D and ε are not necessarily the same. However, FELSENSTEIN (1965) and KIMURA (1965) showed that, in the absence of a modifier gene, the sign of D is quickly adjusted to be the same as that of ε under continuous selection. Therefore, it can be said that the frequency of a recombination-reducing gene generally increases if the loci of which the recombination value is to be modified show epistasis. In other words, for linkage intensity to be modified there must be epistasis, as in Model 1. Note also that, if $r_2 \geqslant r_1 \geqslant r_0$, $r_m - r_M$ and consequently Δx in (9) become negative, so that a modifier gene which enhances recombination can not increase in frequency. This implies that natural selection never increases recombination, if the fitnesses of genotypes remain unchanged.

KIMURA (1965) further showed that, in the absence of a modifier gene, if the population reaches a *quasi linkage equilibrium*, i.e. the state where $\Delta\log(P_1P_4/P_2P_3) = 0$ or $P_1P_4/P_2P_3 = 1 + \varepsilon/r_0$ approximately, the epistatic variance is given by $V_{EP} = r_0\overline{W}D$. Note that this state is quickly attained if ε is much smaller than r_0. Now suppose that in course of time a new mutation occurs at the modifier locus from m to M and its frequency x starts to increase. In the initial stage of the increase where x is small, $r_m - r_M$ is approximately equal to cr_0, where c is the proportion of recombination value reduction in heterozygotes, i.e. $(r_0 - r_1)/r_0$. Therefore, the rate of increase of x is approximately given by

$$\Delta x = xy(cV_{EP})/\overline{W}^2 \tag{10}$$

which further approximates to

$$xy(cV_{EP})/\overline{W} \tag{10'}$$

if selection is mild and $\overline{W}$ is close to 1. The rate of increase of x is obtained also by using a continuous time model and Malthusian fitness parameters instead of Darwinian fitnesses as used here (cf. KIMURA 1956; FELSENSTEIN 1965). It becomes

$$\frac{dx}{dn} = xy(cV_{EP}) \tag{11}$$

Formula (10') or (11) shows an important property that in the initial progress of the modifier gene the selective difference between the modifier and wild-type genes is equal to *the proportion of recombination value reduction in heterozygotes multiplied by the epistatic variance of the loci of which the recombination value is to be modified.* It is of interest to see that this property is similar to FISHER's (1930) fundamental theorem of natural selection, i.e. the rate of increase

of fitness is equal to the additive genetic variance at that time. When r_0, r_1, and r_2 are all large compared with $|\varepsilon|$ and a quasi linkage equilibrium holds, the rate of change of x in the whole process of the modification of recombination value may be given by (10) or (11), replacing $c = (r_0 - r_1)/r_0$ by $c' = (r_m - r_M)/\bar{r}$.

In some specific cases formula (9) or (10) reduces to a simpler form. As an example, consider the case where $W_1 = W_4 = 1 - s$ and $W_2 = W_3 = 1$ $(1 \geqslant s > 0)$. In the absence of modifier gene the equilibrium genotype frequencies of *AB*, *Ab*, *aB*, and *ab* are given by

$$\hat{P}_1 = \hat{P}_4 = \tfrac{1}{4}\left[1 + (\alpha/2) + \alpha\beta - \sqrt{1 + \alpha^2(\tfrac{1}{4} + \beta^2) + \alpha(1 + \alpha\beta)}\right] \quad (12a)$$

$$\hat{P}_2 = \hat{P}_3 = \tfrac{1}{2} - \hat{P}_1 \quad (12b)$$

where $\alpha = r/(1 - r)$ and $\beta = (1 - s)/s$. This equilibrium is stable only if $P_2 = P_3$. JAIN and ALLARD (1966) called this type of equilibrium *isoplethic*. (FINNEY 1952) used this word for a different meaning.) In this connection it should be noted that in the haploid system there is no mechanism for maintaining a strictly stable polymorphism except frequency-dependent selection or alternation of selective advantage between *AB* and *ab* due to environmental changes in different generations. If we superimpose one of these two factors into the present model, then the equilibrium would become stable.

Now suppose that a new mutation arises which reduces the recombination value. After some stochastic changes, this gene is expected to increase in frequency and eventually replace its original type gene in an infinite population. During this replacement the genotype frequencies of *AB*, *Ab*, *aB*, and *ab* change from the values for $\alpha = r_0/(1 - r_0)$ to those for $\alpha = r_2/(1 - r_2)$ in (12 a, b). The selective difference between the modifier and original type genes is dependent on x except for the case of $s = 1$ and $r_0 - r_1 = r_1 - r_2$ (additive gene action). For example, if $s = 0.5$, $r_0 = 0.1$, $r_1 = 0.05$, and $r_2 = 0.0$, the selective advantage of modifier gene, which is given by $(r_m - r_M)D\varepsilon$, changes from 0.0113 to 0.0125 as x increases from 0 to 1. If, on the other hand, $r_0 = 0.1$, $r_1 = r_2 = 0.0$, i.e. the modifier gene shows a complete dominance, $r_m - r_M$ decreases as x increases, while D increases very slightly, so that the selective advantage of the modifier gene gradually decreases as time proceeds, namely, from 0.05 to 0. Figure 1 shows the corresponding changes in x and $\bar{r}$. These changes were obtained numerically by using formula (8) and the following expressions for the frequencies of *AB*, *Ab*, *aB* and *ab* in the next generation.

$$P'_1 = (P_1 - \bar{r}D)W_1/\bar{W} \quad (13a)$$

$$P'_2 = (P_2 - \bar{r}D)W_2/\bar{W} \quad (13b)$$

$$P'_3 = (P_3 - \bar{r}D)W_3/\bar{W} \quad (13c)$$

$$P'_4 = (P_4 - \bar{r}D)W_4/\bar{W} \quad (13d)$$

It is seen from Figure 1 that the rates of changes in x and $\bar{r}$ when x is small are larger in the case of complete dominance than in the case of incomplete dominance but gradually become smaller in the former than in the latter.

When $s = 1$, ε is -2, and D is always $-\tfrac{1}{4}$ irrespective of the value of x, provided that the initial values of P_1, P_2, P_3, and P_4 are 0, 0.5, 0.5, and 0, respectively, which are the equilibrium values in the absence of modifier gene. Further, if

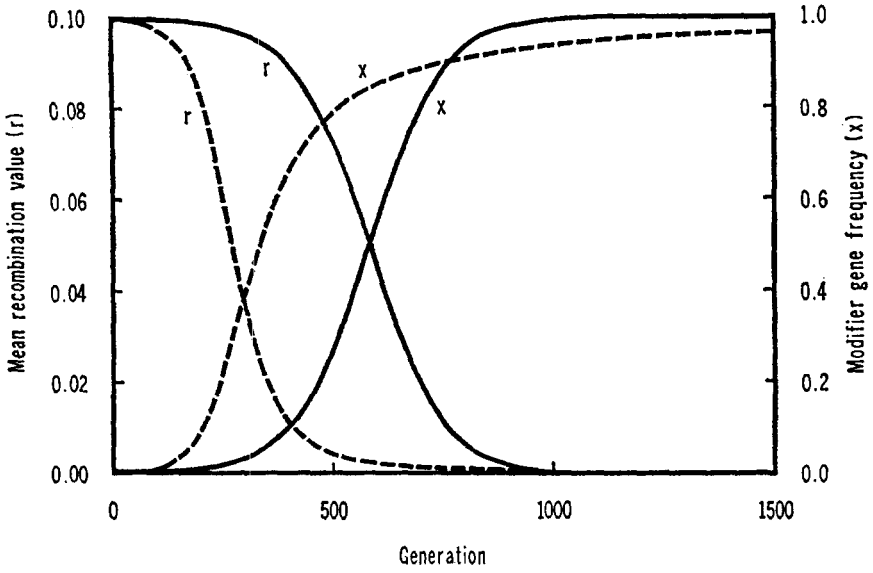

FIGURE 1.—Changes in the modifier gene frequency and the mean recombination value due to natural selection; haploid system. Fitness, $W_1 = 1$, $W_2 = W_3 = 0.5$, $W_4 = 1$. The solid line refers to the case of $r_0 = 0.1$, $r_1 = 0.05$, $r_2 = 0.0$ and the broken line to the case of $r_0 = 0.1$, $r_1 = r_2 = 0.0$.

$r_0 - r_1 = r_1 - r_2 = \Delta r$, the value of $r_m - r_M$ is independent of x and given by Δr. Therefore, the rate of change in x is given by

$$\Delta x = (xy/\overline{W})(\Delta r/2) \tag{14}$$

Since $\overline{W}$ is $1 - r_0/2 + \Delta r x$, the rate of change in x may be represented by

$$\frac{dx}{dn} = \frac{xy\Delta r}{2(1 - r_0/2 + \Delta r x)} \tag{14'}$$

Solution of this equation gives

$$u_n = u_0 + (\Delta r/2)n \tag{15}$$

where

$$u = [1 - (r_0/2)] \log_e [x/(1-x)] - \Delta r \log_e (1-x).$$

Thus, if Δr is large, the frequency of modifier gene increases fairly rapidly. For example, if $r_0 = 0.1$, $r_1 = 0.05$, and $r_2 = 0$ with $\Delta r = 0.05$, x increases from 0.001 to 0.999 in about 540 generations. The mean recombination value correspondingly decreases from 0.0999 to 0.0001. Therefore, the modification of linkage intensity is much more rapid in this case than in the case of $s = 0.5$ (cf. Figure 1). If the modifier gene shows a complete dominance over the original gene, i.e. $r_1 = r_2$, the quantity $r_m - r_M$ is no longer independent of x, and it is difficult to obtain an analytical solution. However, a numerical computation has shown that the changes in x and r are of the same pattern as those for $s = 0.5$, though the modification of recombination value is again more rapid in this case.

In the above examples, the A and B loci remain polymorphic not only in the process of modification of recombination value but also after the completion of this process. Particularly when $s = 1$, the genotype frequencies for these two loci

remain the same during the whole process of modification of recombination value. In all cases, however, the mean fitness for these loci gradually increases as the recombination value is reduced.

The modification of linkage intensity also occurs when the A and B loci show a transient polymorphism or they are going to fixation, if these show epistasis. Consider an example in which $W_1 = 1$, $W_2 = W_3 = 0.5$, and $W_4 = 1.1$. As mentioned previously, this set of fitnesses creates an unstable equilibrium below which the frequency of *ab* cannot increase in the absence of a modifier gene. If the recombination value is allowed to be modified, the increase or decrease of the *ab* frequency depends on the initial values of x, P_1, P_2, P_3, P_4 as well as on the values of r_0, r_1, and r_2. Table 2 shows the results of a numerical computation, in which the initial values of x, P_1, P_2, P_3, and P_4 were 0.01, 0.49, 0.21, 0.21, and 0.09, respectively, with $r_0 = 0.1$, and $r_1 = r_2 = 0$. It is seen that in this case the frequency of *ab* gradually increases and the population eventually becomes monomorphic, consisting of only *ab*. During this process x increases from 0.01 to 0.025, and the mean recombination value is reduced from 0.098 to 0.095. Therefore, the modification of recombination value is very slight. Of course, if x_0 is close to 0.5, the modification is more rapid and greater, though it is never completed unless the fixation of the modifier gene is assisted by random genetic drift. For example, if x starts from 0.5, it increases up to 0.598 before the population becomes monomorphic, while $\bar{r}$ changes from 0.025 to 0.016. Moreover, if r_0 and r_1 are sufficiently large and the gene frequencies of a and b are small, the genotype *ab* is not established in the population but eventually eliminated. In this case the modification of recombination value becomes even smaller.

Diploid system

Model 1: This model corresponds to Model 1 in the haploid system and is essentially the same as KIMURA's (1956). Consider two loci A-a and B-b as before, and let p_1, p_2, p_3, and p_4 be the frequencies after meiosis of the four types of gametes *AB*, *Ab*, *aB*, and *ab* respectively. We denote the fitness of the genotype formed by the union of the ith and jth gametes by W_{ij} as shown in Table 3. Then,

TABLE 2

An example showing the modification of linkage intensity when the A *and* B *loci show a transient polymorphism*

Generation	x	P_1	$P_2(=P_3)$	P_4	D	$\bar{r}$
0	.01000	.49000	.21000	.09000	.00000	.09801
10	.01137	.75096	.01590	.21723	.16288	.09774
20	.01379	.65448	.01917	.30718	.20068	.09726
30	.01734	.48227	.02194	.47385	.22804	.09656
40	.02159	.23212	.01723	.73341	.16994	.09573
60	.02481	.00910	.00105	.98881	.01086	.09510
100	.02493	.00000	.00000	1.00000	.00000	.09508

Fitnesses, $W_1=1$, $W_2=W_3=0.5$, $W_4=1.1$; recombination values, $r_0=0.1$, $r_1=r_2=0.0$.

TABLE 3

Fitnesses of the genotypes for loci A *and* B

	AB	*Ab*	*aB*	*ab*
AB	W_{11}	W_{12}	W_{13}	W_{14}
Ab	W_{21}	W_{22}	W_{23}	W_{24}
aB	W_{31}	W_{32}	W_{33}	W_{34}
ab	W_{41}	W_{42}	W_{43}	W_{44}

$W_{ij}=W_{ji}$.

the amounts of change in gamete frequencies per generation are given by

$$\Delta p_1 = [p_1(W_{1.} - \overline{W}) - rD_W]/\overline{W}$$
$$\Delta p_2 = [p_2(W_{2.} - \overline{W}) + rD_W]/\overline{W}$$
$$\Delta p_3 = [p_3(W_{3.} - \overline{W}) + rD_W]/\overline{W}$$
$$\Delta p_4 = [p_4(W_{4.} - \overline{W}) - rD_W]/\overline{W}$$

as shown by KIMURA (1956, 1965) and LEWONTIN and KOJIMA (1960), where $W_{i.} = \sum_{j=1}^{4} p_j W_{ij}$, $\overline{W} = \sum_{i=1}^{4} p_i W_{i.}$, and $D_W = p_1p_4W_{14} - p_2p_3W_{23}$.

Now suppose that a new completely linked gene pair $\hat{ab}$ is formed from ab and let the frequencies of $\hat{ab}$ and ab be x and p_4^u, respectively, with $x + p_4^u = p_4$. The amounts of changes in p_4^u and x are then given by

$$\Delta p_4^u = [p_4^u\,(W_{4.} - \overline{W}) - rD_W^u]/\overline{W} \qquad (16a)$$
$$\Delta x = x(W_{4.} - \overline{W})/\overline{W} \qquad (16b)$$

where no crossing over is assumed to occur in such genotypes as $AB/\hat{ab}$ and $ab/\hat{ab}$ and D_W^u denotes $p_1p_4^uW_{14} - p_2p_3W_{23}$. Formula (16b) is equivalent to KIMURA's (1956) formula (8). The initial progress of gamete $\hat{ab}$ is again given by formula (3), replacing W_4 by $\hat{W}_{4.}$, i.e. the equilibrium value of $W_{4.}$. Therefore, this type of modification of linkage intensity is fairly rapid, if $\hat{W}_{4.} - \hat{\overline{W}}$ is large.

Model 2: This model represents the modification of linkage intensity by selection of modifier genes. As before, we consider two loci each with two alleles, *A-a*, and *B-b*, and designate the fitnesses of genotypes as given in Table 3. Let P_{ij} be the *frequency after selection* of the genotype formed by the union of the ith and jth gametes, namely $P_{ij} = p_ip_jW_{ij}/\overline{W}$. Then, $P_{i.} = \sum_{j=1}^{4} P_{ij}$ is the frequency of the ith type chromosome before meiosis. This will be called *chromosome frequency*. We will designate by x and $y(=1-x)$ the frequencies of modifier gene M and its allelomorphic gene m respectively. Each genotype for loci A and B can be subdivided into three groups according to the genotype of the modifier locus, i.e. MM, Mm, and mm. But this subdivision is irrelevant to the gamete types produced, except for double heterozygotes. We again assume that the modifier locus is independent of both loci A and B. In the computation of gamete frequencies Table 1 can be used if we replace *mating type* by *genotype* and *genotype produced* by *gamete produced*. Thus, $AB \times AB$, $AB \times Ab$, etc. are replaced by AB/AB, AB/Ab, etc., and P_1^2, $2P_1P_2$, etc. by P_{11}, $2P_{12}$, etc. It will be easily

TABLE 4

Gamete frequencies after meiosis

	M	m	Total
AB	$x(P_{1.} - r_M D)$	$y(P_{1.} - r_m D)$	$P_{1.} - \bar{r}D$
Ab	$x(P_{2.} + r_M D)$	$y(P_{2.} + r_m D)$	$P_{2.} + \bar{r}D$
aB	$x(P_{3.} + r_M D)$	$y(P_{3.} + r_m D)$	$P_{3.} + \bar{r}D$
ab	$x(P_{4.} - r_M D)$	$y(P_{4.} - r_m D)$	$P_{4.} - \bar{r}D$

$r_M = xr_2 + yr_1$, $r_m = xr_1 + yr_0$, $\bar{r} = xr_M + yr_m$, $D = P_{14} - P_{23}$.

seen that the frequencies of various gamete types become as given in Table 4. In this table, D denotes $P_{14} - P_{23}$ or $D_W/\overline{W}$.

The mean fitnesses of genotypes MM, Mm, and mm in the next generation are obtained as follows:

$$\begin{aligned}\overline{W}_{MM} &= (P_{1.} - r_M D)^2 W_{11} + 2(P_{1.} - r_M D)(P_{2.} + r_M D) W_{12} \\ &\quad + 2(P_{1.} - r_M D)(P_{3.} + r_M D) W_{13} + 2(P_{1.} - r_M D)(P_{4.} - r_M D) W_{14} \\ &\quad + (P_{2.} + r_M D)^2 W_{22} + 2(P_{2.} + r_M D)(P_{3.} + r_M D) W_{23} \\ &\quad + 2(P_{2.} + r_M D)(P_{4.} - r_M D) W_{24} + (P_{3.} + r_M D)^2 W_{33} \\ &\quad + 2(P_{3.} + r_M D)(P_{4.} - r_M D) W_{34} + (P_{4.} - r_M D)^2 W_{44} \\ &= W_E - 2r_M D\varepsilon_A + r^2{}_M D^2 \varepsilon_D \\ \overline{W}_{Mm} &= W_E - (r_M + r_m) D\varepsilon_A + r_M r_m D^2 \varepsilon_D \\ \overline{W}_{mm} &= W_E - 2r_m D\varepsilon_A + r^2{}_m D^2 \varepsilon_D\end{aligned}$$

where $W_E = \sum_{i=1}^{4} P_{i.} W_{i.}$, $\varepsilon_A = W_{1.} - W_{2.} - W_{3.} + W_{4.}$, and $\varepsilon_D = E_1 - E_2 - E_3 + E_4$ with $W_{i.} = \sum_{j=1}^{4} P_{j.} W_{ij}$ and $E_i = W_{i1} - W_{i2} - W_{i3} + W_{i4}$. The average fitnesses of gene M and m ($\overline{W}_M$ and $\overline{W}_m$) and the total population ($\overline{W}$) then become

$$\begin{aligned}\overline{W}_M &= x\overline{W}_{MM} + y\overline{W}_{Mm} \\ &= W_E - (r_M + \bar{r}) D\varepsilon_A + r_M \bar{r} D^2 \varepsilon_D \qquad (17) \\ \overline{W}_m &= W_E - (r_m + \bar{r}) D\varepsilon_A + r_m \bar{r} D^2 \varepsilon_D \qquad (18) \\ \overline{W} &= W_E - 2\bar{r} D\varepsilon_A + \bar{r}^2 D^2 \varepsilon_D \qquad (19)\end{aligned}$$

Therefore, the gene frequency of M in the next generation is given by $x\overline{W}_M/\overline{W}$, and the amount of change per generation is

$$\Delta x = \frac{xy}{\overline{W}} (r_m - r_M) D(\varepsilon_A - \bar{r} D\varepsilon_D). \qquad (20)$$

The quantity $\varepsilon_A - \bar{r} D\varepsilon_D$ in the above expression is equivalent to KIMURA's (1965) $\bar{\varepsilon}$, and he showed that the sign of D is quickly adjusted to be the same as that of $\bar{\varepsilon}$ and the population again reaches quasi linkage equilibrium when selection is mild. Therefore, the frequency of a modifier gene generally increases if the loci of which the recombination value is to be modified show epistasis. When there is no epistasis, both ε_A and ε_D are 0, so that no modification of linkage intensity is possible.

As shown by KIMURA (1965), the chromosomal epistatic variance in the state of quasi linkage equilibrium is given by $V_{EPC} = 2r_0 \overline{W} D\bar{\varepsilon}$ in the absence of a modifier gene. Thus, in the initial progress of the modifier gene we again have

$$\Delta x \approx \frac{xy}{W}\left(\frac{c}{2}V_{EPC}\right) \tag{21}$$

or

$$\frac{dx}{dn} = xy\left(\frac{c}{2}V_{EPC}\right) \tag{22}$$

where c is the same as before. Therefore, the rule established for the haploid system holds true also for the diploid system with a slight modification, i.e., substituting *the epistatic variance* by *half the epistatic variance.*

There are numerous types of balanced polymorphisms involving epistasis in diploid populations. FISHER (1930) considered a particular type of epistasis which would be favorable for reducing the recombination value. The fitness matrix given in (a) in Table 5 is one of this type of epistasis. The equilibrium gamete frequencies and their stability conditions for this type of episatsis were investigated by KIMURA (1956). The present model corresponds to the case of $s = 0.1$ and $t = 0.2$ in his formula, and the equilibrium frequencies of *AB*, *Ab*, *aB*, and *ab after meiosis* ($\hat{p}_1$, $\hat{p}_2$, $\hat{p}_3$, and $\hat{p}_4$) are given by

$$\hat{p}_1 = \hat{p}_4 = \tfrac{1}{2}\left[\tfrac{1}{2} - (r/s) + \sqrt{\tfrac{1}{4} + (r/s)^2}\right]$$
$$\hat{p}_2 = \hat{p}_3 = \tfrac{1}{2} - \hat{p}_1$$

where r is the recombination value in the absence of a modifier gene. In this case stable equilibria obtain only if $r < 0.0375$. For other values of r, however, the isoplethic equilibria apply if $p_2 = p_3$ or $P_{2.} = P_{3.}$. Note that p_1, p_2, etc. can be transformed into $P_{1.}$, $P_{2.}$, etc. by the formulas in the column 'Total' in Table 4. The selective advantages of a modifier gene in the initial stage of its increase when r_1 is small, $r_0 D(\varepsilon_A - r_0 D\varepsilon_D)$, for various values of r_0 are given in column (a) in Table 6. It is clear that the selective advantage is very small for any value of r_0, so that a large number of generations are required for the complete modification of linkage intensity. A stronger type of epistasis, where s and t in KIMURA's formula get large, creates a larger selective advantage of the modifier gene than the above, but the stability condition becomes more severe (the critical recombination value below which the stable equilibria exist becomes smaller), though the isoplethic equilibria are possible for any degree of epistasis.

The second set of genotype fitnesses (b) in Table 5 represents an intermediate optimum model with heterozygote advantage. This model was first studied by WRIGHT (1952, 1965) in relation to MATHER's (1943) concept on the relational

TABLE 5

Fitness matrices used for computations of the selective advantage of recombination-reducing gene

	(a)			(b)			(c)			(d)		
	AA	*Aa*	*aa*	*AA*	*Aa*	*aa*	*AA*	*Aa*	*aa*	*AA*	*Aa*	*aa*
BB	.9	1.0	.7	.60	.95	1.00	.5000	.5625	.3750	0	0	0
Bb	.8	1.0	.8	.95	1.10	.95	.5000	1.0000	.4375	0	1	0
bb	.7	1.0	.9	1.00	.95	.60	.3750	.3125	.3750	0	0	0
Model due to	(KIMURA)			(WRIGHT)			(LEWONTIN)			(HALDANE)		

TABLE 6

Selective advantages of the recombination-reducing gene in the initial stage of progress

Initial recombination value	(a)	(b)	(c)	(d)
.01	.00022	.00039	.....	.0048
.02	.00036	.00071	.00256	.0092
.05	.00056	.00133	.00491	.0199
.10	.00065	.00178	.00466	.0296
.20	.00068	.00203	.00065	.0187
.50	.00070	.00214	.00020	.0000

a, b, c, and d refer to the fitness matrices in Table 5.

balance of polygenes. Later, BODMER and PARSONS (1962) and PARSONS (1963) stressed that this type of model is particularly favorable for bringing about the relational balance of genes and consequently for reducing recombination values. The selective advantages of recombination-reducing gene for various initial values of r_0 are given in Table 6 (b). It is seen that the selective advantage is larger in this model than in model (a). In this and the following models of epistasis there are stable equilibria of gamete frequencies before and after the modification of recombination value.

The third model (c), which is a case of cumulative overdominance, was numerically investigated by LEWONTIN (1964), and the selective advantages of a modifier gene in this case can be obtained from Table 5 in his paper. Table 6

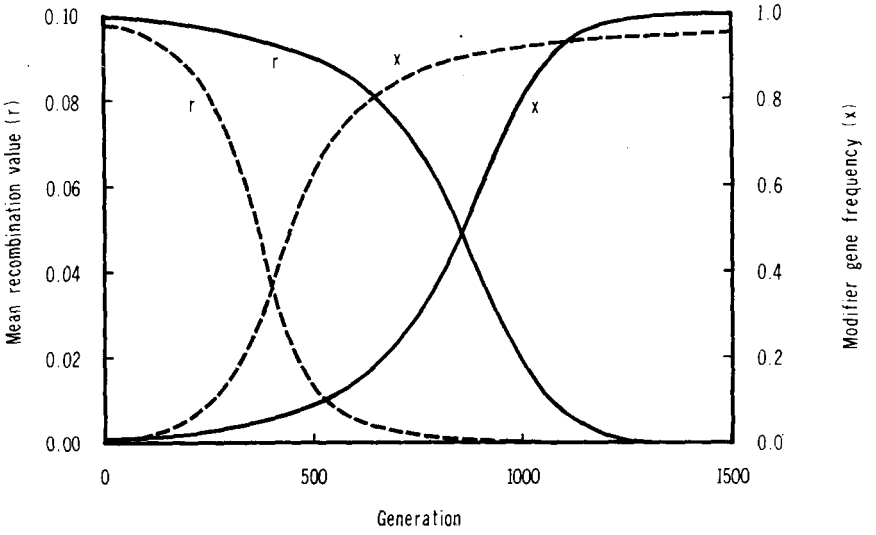

FIGURE 2.—Changes in the modifier gene frequency and the mean recombination value due to natural selection; diploid system. Fitness, $W_{11} = 0.5$, $W_{12} = 0.5625$, $W_{13} = 0.5$, $W_{14} = 1$, $W_{22} = 0.375$, $W_{23} = 1$, $W_{24} = 0.4375$, $W_{33} = 0.3750$, $W_{34} = 0.3125$, $W_{44} = 0.375$. The solid line refers to the case of $r_0 = 0.1$, $r_1 = 0.05$, $r_2 = 0.0$ and the broken line to the case of $r_0 = 0.1$, $r_1 = r_2 = 0.0$.

shows that this model is more favorable than model (b) in modifying the recombination value, although the intensity of selection (genetic load) is also higher in this case. Figure 2 shows the changes in x and $\bar{r}$ in successive generations when $x_0 = 0.01$, $r_0 = 0.1$ and $r_2 = 0.0$.

The last model (d) represents a case of balanced lethals. As shown by HALDANE (1962), the equilibrium chromosome-frequencies in this case are all 0.25 if $r > 1/4$, and if $r < 1/4$ they are

$$\hat{P}_{1.} = \hat{P}_{4.} = 1/2\ [1 \pm (1 - 4r)^{1/2}(1 - r)]$$
$$\hat{P}_{2.} = \hat{P}_{3.} = 1/2 - \hat{P}_{1.}$$

Thus, the initial selective advantage $r_0 D(\varepsilon_A - r_0 D\varepsilon_D)$ is 0 when $r_0 > 1/4$ and $r_0(1 - 2r_0)\ (1 - 4r_0)/2(1 - r_0)^2$ when $r_0 < 1/4$. Table 6 shows that the selective advantage of a modifier gene is very high in this case and thus this type of balanced lethals is expected quickly to reduce the recombination value between them if modifiers exist. A numerical computation has shown that the modifier gene frequency changes from 0.01 to 0.99 in about 190 generations when $r_0 = 0.1$, $r_1 = 0.05$, and $r_2 = 0.0$.

As in the case of haploid system the modification of linkage intensity is possible also when the A and B loci are going to fixation. One example is given in Table 7 where the fitness set represents an optimum model involving no stable equilibria, namely, $W_{11} = W_{44} = 0.6$, $W_{12} = W_{13} = W_{24} = W_{34} = 0.9$, and $W_{14} = W_{22} = W_{23} = W_{33} = 1.0$. The initial values of $P_{1.}$, $P_{2.}$, $P_{3.}$, and $P_{4.}$ employed were 0, 0.5001, 0.4999, and 0, respectively, with $r_0 = 0.1$, $r_1 = 0.0$, and $r_2 = 0.0$. It is seen from Table 7 that in this case the selection of a modifier gene is not so effective as with stable polymorphic loci.

DISCUSSION

Development of linkage: At present it is not known how and when the linkage of genes was developed in the evolutionary history of life. The fact that bacteria and viruses have genetic linkage suggests that it occurred in the very early stage

TABLE 7

An example showing the modification of linkage intensity when the A *and* B *loci show a transient polymorphism*

Generation	x	$P_{1.}$	$P_{2.}$	$P_{3.}$	$P_{4.}$	D	$\bar{r}$
0	.01000	.00000	.50010	.49990	.00000	—.25000	.09801
10	.01038	.10981	.39031	.39007	.10981	—.14177	.09794
50	.01142	.13383	.36656	.36578	.13383	—.11173	.09773
80	.01224	.13375	.36720	.36529	.13375	—.11180	.09757
100	.01282	.13367	.36808	.36458	.13367	—.11188	.09746
200	.01615	.13265	.40311	.33160	.13265	—.11174	.09680
400	.01945	.00002	.99996	.00000	.00002	—.00000	.09615

Fitness, $W_{11} = W_{44} = 0.6$, $W_{12} = W_{13} = W_{24} = W_{34} = 0.9$, and $W_{14} = W_{22} = W_{23} = W_{33} = 1.0$; recombination values, $r_0 = 0.1$, $r_1 = r_2 = 0$.

of organic evolution. The linkage of genes may occur in two different ways. One is that a gene (or genes) is first duplicated, resulting in an end-to-end connection of the duplicated and original genes (tandem duplication), and in course of time the two genes are gradually differentiated owing to mutation so as to have different biochemical or biological functions. In this case linkage is made first and the recombination mechanism is developed later. In the second way, the duplicated gene is not connected with the original gene from the beginning but exists independently of the latter, and in course of time some type of epistasis or gene interaction arises and then the linkage of genes, which may occur by some mechanism, becomes selectively advantageous and increases in frequency. In this case the duplicated gene may not necessarily have originated as duplicated gene of its partner, but may be an already differentiated gene derived from another cell by means of an unequal recombination. Anyway, this case is represented by Model 1 in the section *Haploid system*, where $r = 1/2$.

However, the advantage of linkage in the early stage of organic evolution may not always be due to epistasis, but linkage *per se* may be advantageous. For example, linked genes might be physicochemically more stable than unlinked genes, or the replication of DNA or RNA molecule as gene material might be easier when genes are linked than when they exist independently. In the replicon hypothesis of JACOB, BRENNER and CUZIN (1963) the replication of DNA molecule occurs as an unit of many genes (cistrons). Further, in the first model of the development of linkage in the above paragraph, the duplicated gene pair might be selectively advantageous over the single gene owing to increased formation of an enzyme. In these cases the change in the frequency of linked gene pair $\hat{AB}$ or $\hat{AA}$ (x) may be described by

$$\frac{dx}{dn} \approx \frac{x(1-x)s}{1-s(1-x)}$$

where s is the selective advantage of $\hat{AB}$ or $\hat{AA}$ over the unlinked gene pair AB or single gene A. In these cases, therefore, the development of linkage is fairly rapid unless s is very small.

Modification of linkage intensity: In the treatment of modification of linkage intensity in Model 2, we assumed the existence of modifier genes. As mentioned previously, there are a large number of investigations which indicate that recombinations between gene loci are under genetic control. The recombination-modifying genes so far identified presumably control the formation of some enzyme(s) involved in the recombination event, and the physical distance between gene loci would not be affected in this case. The positive interference between recombination events generally observed in higher organisms, or the lack of recombination in male Drosophila perhaps reflects this type of recombination control. The fact that environmental and genetic factors affect crossing over mainly in the centromere region in Drosophila supports this concept (see BODMER and PARSONS 1962). The same mechanism might be involved in the influence of an inversion chromosome on crossing over in other chromosomes as observed by LEVINE and LEVINE (1954) in *Drosophila pseudoobscura*. Presumably, the formation of recombinational enzyme(s) is affected by inversion through a position effect.

There is, however, another type of recombination control, namely physical changes of gene arrangement of chromosome structure. Translocations and inversions are typical examples. In this case the changes of recombination values are more direct. However, chromosome rearrangements may sometimes destroy a gene complex or operon already established, resulting in a so-called position effect (cf. LEWIS 1967), so that they are not always selectively advantageous. It seems that the real control of recombination values has been mediated both by selection of modifier genes and by chromosome rearrangements. If the degree of epistasis is the same, the modification of recombination values is more rapid in the latter process than in the former. However, even in the former process the modification of linkage intensity may proceed more rapidly than the modification of dominance as visualized by FISHER (1928).

So far, gene complexes, gene clusters, or operons appear to be more common in bacteria and bacteriophages than in fungi and other higher organisms. HOROWITZ (1965) and LEWIS (1967) postulate that operons have often arisen by a process of repeated tandem duplications accompanied by gradual functional differentiation of the daughter genes, and operons for the primitive biosynthetic pathways of living organisms seem to have become fragmented in higher organisms. The tight linkage between the β and δ loci of human hemoglobin (BOYER *et al.* 1963), for example, advocates this hypothesis for the origin of operons or gene complexes. If, however, operons or gene complexes are a unit of strong gene interaction as shown by STAHL and MURRAY (1966) in the T2 and T4 phages, the gradual tightening of linkage between structural genes would also have been possible, as shown in this paper. If organisms live in a relatively constant environment for a long time with little evolution, as bacteria and bacteriophages might have done, this possibility cannot be ruled out. In organisms which have evolved rapidly, recombinations would have been necessary for combining various advantageous mutations into the same individual, as indicated by FISHER (1930), MULLER (1932), and CROW and KIMURA (1965). Therefore, from this point of view too, the scarcity of operons or gene complexes in higher organisms is expected. Further, even if gene complexes had originated as tandem duplications, gene interactions would have played an important role in maintaining them in populations for a long time.

SUMMARY

Two different mathematical models for the modification of linkage intensity by natural selection are presented. In the first model the changes of recombination values are assumed to occur by structural changes of chromosome such as translocation and inversion, while in the second they are brought about by selection of modifier genes. In both models there must be epistasis or gene interaction for the modification of linkage intensity to proceed, the speed of the modification depending upon the degree of epistasis, and this modification always occurs in the direction of decreased recombination, if the fitnesses of genotypes remain constant. If the degree of epistasis is the same, the modification of linkage intensity is more rapid in the first model than in the second. In general, stable polymor-

phisms with epistasis are more favorable for establishing close linkage than unstable and transient polymorphisms with epistasis. Two different models for the evolution of linkage are proposed.

LITERATURE CITED

BARRY, E. G., 1967 Chromosome aberrations in Neurospora, and the correlation of chromosomes and linkage groups. Genetics **55**: 21–32.

BECKWITH, J. R., and E. R. SIGNER, 1966 Transposition of the *lac* region of *Escherichia coli*. I. Inversion of the *lac* operon and transduction of *lac* by ϕ80.. J. Mol. Biol. **19**: 254–265.

BODMER, W. F., and P. A. PARSONS, 1962 Linkage and recombination in evolution. Advan. Genet. **11**: 1–100.

BOYER, S. H., D. L. RUCKNAGEL, D. J. WEATHERALL, and E. J. WATSON-WILLIAMS, 1963 Further evidence for linkage between the β and δ loci governing human hemoglobin and the population dynamics of linked genes. Am. J. Human Genet. **15**: 438–448.

CLARK, A. J., and A. D. MARGULIES, 1965 Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12. Proc. Natl. Acad. Sci. U.S. **53**: 451–459.

CROW, J. F., and M. KIMURA, 1965 Evolution in sexual and asexual populations. Am. Naturalist **99**: 439–450.

DETLEFSEN, J. A., and E. ROBERTS, 1921 Studies on crossing over. I. The effect of selection on crossover values. J. Exptl. Zool. **32**: 333–354.

FELSENSTEIN, J., 1965 The effect of linkage on directional selection. Genetics **52**: 349–363.

FINNEY, D. J., 1952 The equilibrium of a self-incompatible polymorphic species. Genetica **26**: 33–64.

FISHER, R. A., 1928 The possible modification of the response of the wild type to recurrent mutants. Am. Naturalist **62**: 115–126. —— 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

HALDANE, J. B. S., 1962 The selection of double heterozygotes. J. Genet. **58**: 125–128.

HOROWITZ, N. H., 1965 The evolution of biochemical syntheses—retrospect and prospect. pp. 15–23. *Evolving Genes and Proteins*. Edited by V. BRYSON and H. J. VOGEL. Academic Press, New York.

HOWARD-FLANDERS, P., and L. THERIOT, 1966 Mutants of *Escherichia coli* K-12 defective in DNA repair and in genetic recombination. Genetics **53**: 1137–1150.

JACOB, F., and E. L. WOLLMAN, 1961 *Sexuality and the Genetics of Bacteria*. Academic Press, New York.

JACOB, F., S. BRENNER, and F. CUZIN, 1963 On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp. Quant. Biol. **28**: 329–348.

JAIN, S. K., and R. W. ALLARD, 1966 The effects of linkage, epistasis, and inbreeding on population changes under selection. Genetics **53**: 633–659.

KIKKAWA, H., 1937 Spontaneous crossing over in the male of *Drosophila ananassae*. Zool. Mag. (Tokyo) **49**: 159–160.

KIMURA, M., 1956 A model of a genetic system which leads to closer linkage by natural selection. Evolution **10**: 278–287. —— 1965 Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics **52**: 875–890.

LEVINE, R. P., and E. E. LEVINE, 1954 The genotypic control of crossing over in *Drosophila pseudoobscura*. Genetics **39**: 677–691.

LEWIS, E. B., 1967 Genes and gene complexes. pp. 17–47. *Heritage from Mendel*. Edited by R. A. BRINK. Univ. Wisconsin Press, Madison.

LEWONTIN, R. C., 1964 The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49**: 49–67.

LEWONTIN, R. C., and K. KOJIMA, 1960 The evolutionary dynamics of complex polymorphisms. Evolution **14**: 458–472.

MATHER, K., 1943 Polygenic inheritance and natural selection. Biol. Rev. **18**: 32–64.

MORIWAKI, D., 1940 Enhanced crossing over in the second chromosome of *Drosophila ananassae*. Japan. J. Genet. **16**: 37–48.

MUKHERJEE, A. S., 1961 Effect of selection on crossing over in the males of *Drosophila ananassae*. Am. Naturalist **95**: 57–59.

MULLER, H. J., 1932 Some genetic aspects of sex. Am. Naturalist **64**: 118–138.

PARSONS, P. A., 1958 Selection for increased recombination in *Drosophila melanogaster*. Am. Naturalist **92**: 255–256. —— 1963 Polymorphism and the balanced polygenic combination. Evolution **17**: 564–574.

STAHL, F. W., and N. E. MURRAY, 1966 The evolution of gene clusters and genetic circularity in microorganisms. Genetics **53**: 569–576.

WRIGHT, S., 1952 The genetics of quantitative variability. pp. 5–41. *Quantitative Inheritance*. Her Majesty's Stationery Office, London. —— 1965 Factor interaction and linkage in evolution. Proc. Roy. Soc. Lond. B **162**: 80–104.