

THE INTERACTION OF SELECTION AND LINKAGE. II. OPTIMUM MODELS¹

R. C. LEWONTIN²

Department of Biology, University of Rochester, Rochester, N.Y.

Received June 16, 1964

IN the first paper in this series (LEWONTIN 1964) I discussed the consequences of linkage in populations subject to natural selection when there is *heterosis* (higher fitness of heterozygote than of homozygotes). There is a great deal of theoretical and experimental evidence concerning the existence of heterosis at the level of single loci (so-called *overdominance*), but it seems to be as yet undetermined whether such overdominance is widespread or only of academic interest.

On the other hand, there can be no doubt at all that "optimizing selection" must be an extremely common and important biological phenomenon. By "optimizing selection" I mean that in a continuum of possible phenotypes, there is some intermediate value which has the greatest selective advantage and that deviations on either side of this intermediate phenotype are less fit on the average. It is a commonplace that there is a relatively narrow range of optimal size for an animal and although there is considerable genetic variation for size in both directions in a population of say, *Drosophila*, the population tends to remain at an intermediate phenotypic level. Even a component of fitness like fecundity can be increased over its normal value by artificial selection (KOJIMA and KELLEHER 1963) so that correlation with other fitness components somehow holds fecundity at an intermediate value in natural populations.

Even though the widespread existence of "optimizing selection" cannot be questioned, what is still not thoroughly understood is the extent to which such selection can maintain genetic variation in a population. FISHER (1930) first alluded to an optimum selection scheme later more fully developed by MATHER (1941). FISHER's scheme was that there were two factors *A*, *a* and *B*, *b* such that "*A* is advantageous in the presence of *B* but disadvantageous in the presence of *b*, and that *B* is advantageous in the presence of *A* but disadvantageous in the presence of *a*." Couched, as this description is, in terms of gene advantages rather than diploid genotypic fitnesses, it is not unambiguous in its interpretation and as it stands is neither a sufficient nor necessary description of the kind of two factor interaction necessary for stable equilibrium of gene frequencies and for linkage effects of a permanent sort. In the first careful treatment of a two factor polymorphism, KIMURA (1956) showed, in fact, that in order for the system

¹ This investigation was performed under Atomic Energy Commission Contract AT(30-1)2620. The extra cost of setting tables and formulas, and the page charge has been defrayed by this contract.

² Present address: Department of Zoology, University of Chicago, Chicago, Illinois.

described by FISHER to be stable, at least one of the factor pairs would have to be *unconditionally* heterotic.

FISHER went further and considered an explicitly optimum model in his discussion of metrical characters. Here he assumed that the mean of the population was the optimum and that at a single locus the fitness of a genotype fell off with its squared deviation from the mean. FISHER then went on to show that such a system does not lead to a stable equilibrium but that the gene frequencies might be balanced by recurrent mutation and thus account for the observed genetic variability for metric characters. The absolute instability of this model of FISHER's arises because fitness falls off with deviation from the *population mean* and not from a fixed optimum phenotype.

Another treatment of linkage and optimizing selection and one that in some part grew out of FISHER's earlier suggestion was that of MATHER (1941). MATHER's suggestion of "internal" and "relational balance" amounts to a model in which an intermediate phenotype is most fit for a character determined by a large number of loci of similar small action, "polygenes". At each of these loci there are plus and minus modifiers of the phenotype and no epistasis is assumed on the primary phenotypic scale. In general MATHER does not specify how fitness falls off with deviation from the optimum phenotype and the only numerical case he presents of the model (on page 186 of his 1941 paper) does not lead to a stable equilibrium of gene frequencies. Like FISHER, MATHER predicts the buildup of linked repulsion complexes of genes as a result of the optimizing selection, but the models explicitly discussed are not stable. BODMER and PARSONS (1962) present several models based upon the relational balance theory of MATHER, models that do lead to the predictions made by MATHER. However, all such models have, in addition to the optimality of intermediate phenotypes, an assumption of superior heterozygote fitness not obviously related to the phenotypic score. Thus, it is assumed by BODMER and PARSONS that the double heterozygote $Aa Bb$ is more fit than the so called "balanced homozygotes" $AA bb$ and $aa BB$. More recently PARSONS (1963) tried to relax this assumption of heterozygote superiority but the stability, in fact, disappears with the heterosis. In this connection the point made by KIMURA (1956) must be reiterated, that interaction and linkage alone cannot maintain stable gene frequency equilibria and linkage complexes. There must be some heterosis of fitness as well. Thus the balance models of BODMER and PARSONS can be subsumed under the general rule that epistatic deviations of any kind will lead to the buildup of stable linked complexes when recombination is restricted, *provided that each locus shows marginal overdominance for fitness at equilibrium* (KOJIMA 1959a).

WRIGHT (1935) considered a *quadratic deviation* model in which the fitness of a phenotype falls off as the square of the deviation of that phenotype from some optimum. WRIGHT showed that if the genes controlling the phenotype have either complete dominance or complete additivity, there could be no stable equilibrium of gene frequencies. Eventually all genes controlling the character would be fixed. Essentially the same conclusion was reached by ROBERTSON (1956) for a greater variety of optimum models. Like WRIGHT, ROBERTSON

examined only those cases in which phenotype was determined by additive or completely dominant genes. ROBERTSON's optimum models included selection at the population mean, selection at an intermediate fixed value and selection proportional to an exponential function of the squared deviation from the optimum. Again, no stable equilibria of gene frequencies was predicted. These two studies, then, seemed to rule out optimizing selection as a force maintaining genetic variation.

A discovery of importance in this matter was made by KOJIMA (1959b). He found that there *can* be stable gene frequency equilibria with WRIGHT's quadratic deviation model for more than one locus. The key point is that although complete dominance and complete additivity do not lead to stability, *partial* dominance can lead to such stability provided the degree of dominance falls within a range dependent upon the value of the optimum. This discovery of KOJIMA then revived the optimum model as a possible source of genetic variation.

In none of the cases discussed by these three last authors has the question of linkage arisen. In fact, linkage is likely to play an important role in optimizing selection because optimum models generate a very large amount of epistasis on the fitness scale, much larger than other kinds of selection, and it has been shown by LEWONTIN and KOJIMA (1960) BODMER and PARSONS (1962) and LEWONTIN (1964) that large amounts of epistasis lead to large linkage effects. In fact as this paper will show, even when no permanent stable equilibrium of gene frequencies is predicted in an optimum model with free recombination, restriction of recombination may result in nearly permanent maintenance of genetic variation.

The plan of this paper is to examine a number of different optimum models including WRIGHT's quadratic deviation model to see in what way restriction of recombination affects the genetic structure of the population.

The Quadratic Deviations Model Without Linkage

Suppose that there are a number of loci controlling some partial phenotype of an organism. We assume that each of these loci has two alleles B_j and b_j and that the phenotypic effects, S_j of the three zygotic types are:

$$\begin{array}{ccc} B_j B_j & B_j b_j & b_j b_j \\ a_j & h_j a_j & -a_j \end{array}$$

When $h = 1$ there is complete dominance of B_j , and when $h = 0$ there is no dominance of either allele. Then in a random mating population the contribution of the j th locus to the mean phenotype is

$$(1) \quad P_j = a_j q_j^2 + 2q_j P_j a_j h_j - a_j P_j^2$$

If the loci are additive in their effect on the phenotype, the phenotypic mean of the population, \bar{S} is simply

$$(2) \quad \bar{S} = \sum_{j=1}^n P_j$$

We further suppose that the fitness of any individual declines as the square of the individual's deviation from some optimum phenotype, \bar{O} . That is.

$$(3) \quad W_i = K - (S_i - \bar{O})^2$$

where K is an arbitrary constant.

Then the mean fitness of the population, \bar{W} , is given by

$$(4) \quad \bar{W} = K - [(\bar{S} - \bar{O})^2 + V]$$

where V , the total genotypic variance on the phenotype scale is

$$(5) \quad V = \sum_{j=1}^n [q_j^2 a_j^2 + 2p_j q_j h_j^2 a_j^2 + p_j^2 a_j^2 - P_j^2]$$

when all the loci are in linkage equilibrium.

KOJIMA (1959) showed that stable equilibria of the gene frequencies, q_j exist for this model provided the following three conditions are met:

$$(6a) \quad \frac{d\bar{W}}{dq_j} = 0$$

$$(6b) \quad \frac{d^2\bar{W}}{dq_j^2} < 0$$

$$(6c) \quad \frac{d^2\bar{W}}{dq_i^2} \cdot \frac{d^2\bar{W}}{dq_j^2} > \left[\frac{d^2\bar{W}}{dq_i dq_j} \right]^2$$

for all i and j . When written *in extenso* these derivatives are:

$$(7a) \quad \frac{d\bar{W}}{dq_j} = -2[a_j^2(h_j^2-1)(1-2q_j) + 2a_j\{1+h_j(1-2q_j)\}(\bar{S}-P_j-\bar{O})]$$

$$(7b) \quad \frac{d^2\bar{W}}{dq_j^2} = 2[2a_j^2(h_j^2-1) + 4h_j a_j(\bar{S}-P_j-\bar{O})]$$

$$(7c) \quad \frac{d^2\bar{W}}{dq_i dq_j} = -8 a_i a_j [1+h_j(1-2q_j)][1+h_i(1-2q_i)]$$

Condition 6a simply defines the value of q_j at equilibrium and is equivalent to requiring that the additive genetic variance of fitness is zero. Condition 6b is a requirement for marginal heterosis at each locus at equilibrium and condition 6c is a requirement on the so called "additive by additive" epistasis. It should be noted that this epistasis is on the *fitness scale* and not on the primary phenotypic scale which is assumed to be additive between loci. This epistasis arises from the relation between phenotype and fitness expressed in equation 3. Moreover, no matter how many loci are involved in the character there is only the two-locus epistatic interaction expressed in equation 7c. No other kinds of epistasis are generated by the quadratic optimum model.

By considerable numerical computation using relations 6a-6c KOJIMA was able to map out the possible stable equilibrium situations for two loci on the assumption that the values of a and h are equal for both loci. The result of those computations is shown in Figure 1, modified from KOJIMA's paper to generalize his results for any number of loci. Along the ordinate are values of the optimum, \bar{O} , scaled in units of a . Along the abscissa are the various values of dominance, h . The shaded region represents the combinations of \bar{O} and h that will lead to stable intermediate frequencies of q_1 and q_2 such that both loci are segregating. Outside this area there is no stable equilibrium except the trivial ones of one or both loci fixed. We will return to this point later.

Two points not brought out by KOJIMA need to be made for our investigation.

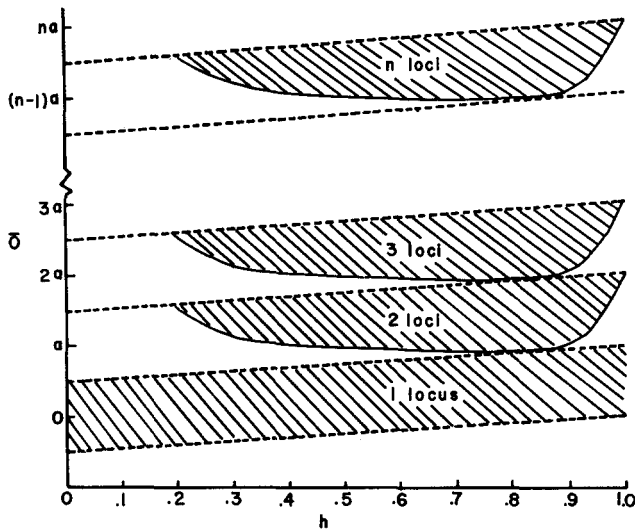


FIGURE 1.—A generalization of KOJIMA'S requirements for stability of the quadratic optimum mode. The ordinate is the value of the optimum \bar{O} scaled in units of gene effect, a . The abscissa is the dominance h . The shaded areas are the regions of stability for successively larger numbers of loci. The dashed lines enclose the region of necessary stability.

First, for a fixed value of h , the equilibrium gene frequencies get closer and closer to unity as \bar{O} is increased. At the upper boundary of the shaded region those frequencies reach unity. At the lower boundary the gene frequencies are closer to .50 but do not reach it and are in fact bounded at much higher frequencies than one half. For example at $h = .8$ when the area has its maximum downward extension, $q = .73$ for both loci. Thus equilibrium frequencies are strongly biased toward high values. Second, the shaded region is bounded by a larger region between the two dashed parallel lines, forming a band of width a . This band is a *necessary* condition on \bar{O} for stability for each value of h . This necessary condition is given by Equations 6a and 6b and simply requires overdominance at each locus at equilibrium. The smaller shaded area gives the *necessary* and *sufficient* condition for stability and is smaller because of the extra requirement on \bar{O} imposed by Equation 6c specifying the amount of epistasis between the loci. For the purposes of our present discussion we need only consider the weaker condition given by the parallel dotted lines, since it will be quite strong enough to make the point being aimed at.

Figure 1 shows a series of such bands enclosed by parallel lines and within each band a shaded area. Each successive band represents the requirements for stable equilibrium of successively larger numbers of loci, each with an additive effect a and a dominance h . The proof of this relationship is given in Appendix I of this paper. What the figure shows is that *to maintain successively larger numbers of loci in stable equilibrium, the optimum phenotype must be successively a larger and larger proportion of the extreme phenotype*. In fact Figure 1 shows

that to maintain n loci in an unfixed state where all loci are of equal additive effect, a , requires that the optimum phenotype fall somewhere between na and $(n-3/2)a$. That is, the absolute difference between optimum and extreme cannot be greater than $3/2 a$. Thus, to maintain, say, ten loci in stable unfixed equilibrium would require that the optimum phenotype be within 15 percent of the extreme possible for these ten loci.

This may be looked at in another way. Suppose 20 loci each with an additive effect of 3 and a dominance of .5 are segregating. The maximum phenotype due to segregating loci is then 60. If the optimum phenotype is 28, all 20 loci cannot be kept in equilibrium. Loci will be fixed one by one until the maximum effect of segregating loci is less than 30 and this will happen when ten loci are segregating. The ten fixed loci will have been fixed half at the plus allele and half at the minus allele so that the mean phenotype is not changed.

The unfixed loci will be maintained in stable equilibrium at gene frequencies very close to unity. However, this is not the only possible stable configuration. If one of the loci should become fixed by chance at the plus allele, the system decays one step to the next lower stability band. That is because any loci fixed at the plus allele can be regarded as scale shifts in Figure 1, shifting both the phenotypes and the value of the optimum, but preserving the difference between them. Each time a locus is fixed the remaining loci equilibrate at gene frequencies closer to .50 with the result that each successive random fixation becomes less likely. The number of loci segregating at equilibrium will be a function of recurrent mutation rates and population size and will increase with an increase in both of these factors (KIMURA and CROW 1964). The fewer loci that are maintained at equilibrium the farther the gene frequency at each locus is from unity for a given position inside the shaded area. For example with $a = 1$, $h = .8$ and optimum one less than the number of loci, two loci can be maintained each at a frequency of .771 while 100 loci will be maintained each at a frequency of .982. Thus, the more loci maintained at equilibrium, the smaller the genetic variance per locus. The interesting question then arises as to how the total genetic variance on the phenotypic scale changes as the number of identical loci increases. That is, how much genetic variance can be maintained in a population by quadratic optimum selection?

TABLE 1

Equilibrium gene frequency and genetic variance of a character for which $h = .8$, $a = 1$ and $0 = (n - 1)$ when different numbers of genes are segregating

n	q	σ^2/locus	Total σ^2
1	.500	.660	.660
2	.771	.193	.386
5	.868	.072	.362
10	.911	.036	.358
100	.982	.003	.259
∞	~ 1	~ 0	.180

Table 1 shows the calculations based on equations 1 and 5 of the phenotypic variance maintained at equilibrium for different numbers of loci when $a = 1$, $h = .8$ and $\bar{O} = (n-1)$.

The table shows that the variance per locus goes down more quickly than the number of loci goes up so that the *total variance actually decreases with increasing number of loci*. This remarkable result then enables us to predict that as more loci concerned with the character are fixed by chance, with fewer and fewer loci segregating at equilibrium, the phenotypic variance for the character will actually *increase*, reaching a maximum when only a single locus is still segregating with a gene frequency close to .50.

The minimum variance as the number of loci grows indefinitely large is

$$(8) \quad \lim_{n \rightarrow \infty} V = a^2(1-h) \left(\frac{1+h}{2} - E \right)$$

where aE is the *absolute* difference between the optimum for an n -locus system and $(n-1)a$. This is a constant irrespective of n since the various bands, as already pointed out, are rigid translations. Thus, for the case we have been considering where $h = .8$, $a = 1$ and $\bar{O} = (n-1)$, the value of E is 0 and

$$\lim_{n \rightarrow \infty} V = \frac{a^2(1-h^2)}{2} = .180$$

as shown in Table 1. The proof of expression 8 is given in Appendix II.

When we turn to the *maximum* variance that can be maintained as the result of a single segregating locus, we can use the usual expression for equilibrium at a single locus

$$(9) \quad \hat{q} = \frac{W_{bb} - W_{Bb}}{W_{bb} + W_{BB} - 2W_{Bb}}$$

Substituting (3) into (9) and using the scale of a , ah and $-a$ for the phenotype we get

$$(10) \quad \hat{q} = \frac{1}{2} + \frac{\bar{O}}{a(1-h^2) + 2h\bar{O}} = \frac{1}{2} + x$$

Moreover, the total phenotypic variance from (5) and (10) is

$$(11) \quad V = \left(\frac{1}{2} - 2x^2 \right) a^2 \left[1 + \frac{h^2}{2} - 2xh(2-xh) \right]$$

For x and h of the same sign, which is the only case of interest to us, the maximum value of (11) occurs when the optimum is exactly halfway between the two homozygotes ($x = 0$) and the variance is simply

$$(12) \quad V = \frac{a^2}{2} \left(1 + \frac{h^2}{2} \right)$$

whose maximum value tends to $.75 a^2$ as h tends to unity. For $\bar{O} = 0$ to result in a stable equilibrium, h must be less than unity. Thus the maximum variance that can be maintained by genes selected in a quadratic optimum selection system is *three quarters of the squared gene effect*. This occurs when the optimum is intermediate between the two homozygotes at a single locus and when the degree of dominance is nearly but not quite unity. Without any dominance ($h = 0$), the variance is half the squared gene effect.

Our conclusion must be, then, that selection based upon the squared deviations from an optimum cannot maintain much variance for a character although it may maintain large numbers of loci segregating. However, when large numbers of loci are segregating each is maintained so close to fixation that random events are sure to reduce the number of segregating loci to very few where the net selection per locus becomes more substantial.

Stable equilibria are also possible when $h > 1$ (overdominance) and for such cases the necessary condition for equilibrium is simply

$$(13) \quad \bar{O} > \left(n - \frac{1}{2}\right)a + \frac{ha}{2}$$

For such overdominant cases the closer the optimum is to its lower limit, given by (13), the closer the equilibrium gene frequencies are to fixation. This is the opposite of the cases of partial dominance where optima close to the lower allowable limit gave gene frequencies farthest from fixation. The reason is that for overdominant cases, large optima are always closer to the heterozygote than the homozygotes, and always closer to the larger homozygote. However, the larger the optimum the more nearly equal are the *relative* deviations of the two homozygotes and therefore the more nearly alike their fitnesses.

The Effect of Linkage on the Quadratic Model

To test whether linkage makes a substantial difference to the conclusions of the previous section I have examined a number of 2-locus and 5-locus quadratic deviation cases numerically. The numerical methods are the same as those used in the first paper of this series, that is the method of *genetic operators* using a digital computer.

Tables 2a and 2b give the pertinent parameters for two 2-locus models. At the top of each table are given the values of a , h and \bar{O} , the assumption being that the two loci are identical in their action. In the body of each table are given the fitnesses of the nine genotypes calculated from the parameters and from relation (3) above. For Model I, $K = 334.0$ and for Model II, $K = 494$. The fitnesses are then adjusted to make the maximum fitness unity. These models satisfy the necessary and sufficient conditions for gene frequency equilibrium given in Figure 1. When $a = 6$ and $h = .6$, the optimum must lie between 5.5 and 10.8 and Models I and II have been chosen to lie just within this interval. These cases

TABLE 2
Parameters and fitnesses for Models I and II

	a. Model I			b. Model II		
	$a=6$	$h=0.6$	$\bar{O}=6$	$a=6$	$h=0.6$	$\bar{O}=10$
	<i>AA</i>	<i>Aa</i>	<i>aa</i>	<i>AA</i>	<i>Aa</i>	<i>aa</i>
<i>BB</i>	.8960	.9651	.8960	.9923	1.0000	.7979
<i>Bb</i>	.9651	1.0000	.7919	1.0000	.9846	.6889
<i>bb</i>	.8960	.7919	.0301	.7979	.6889	.0203

were chosen to cover extremes of stable gene frequency possible with the quadratic optimum model and also different selection intensities. In neither case are the selection intensities very great at equilibrium since the most frequent genotypes, by far, are the double homozygote $AABB$ and the two single heterozygotes whose fitnesses are nearly the same.

The results for these models are shown in Tables 3 and 4. The tables give the recombination fraction R , the four gametic frequencies, the gene frequencies, p , which are the same for both loci, the linkage disequilibrium parameter D' , and the mean fitness of the population \bar{W} . D' is defined in LEWONTIN (1964) as the linkage disequilibrium relative to the maximum possible value, given the gene frequencies. For Model I there are considerable changes in gene frequencies and gametic frequencies with changing recombination, the chief difference appearing between complete linkage ($R = 0$) and about 20 percent recombination. The differences in gametic frequencies are quite profound. Linkage increases the repulsion gametes Ab and aB until they become 80 percent of the gamete pool at $R = 0$ while for free recombination they represent only 35 percent of the pool. This corresponds to the magnitude and sign of the relative linkage disequilibrium parameter, D' , which is consistently negative indicating an excess of repulsion linkages. Unlike the heterotic models examined in the first paper of this series (LEWONTIN 1964) there are no complementary equilibria with an excess of coupling linkages. As will become apparent in the course of this paper, *optimum models are characterized by an excess of repulsion gametes*. The second point worth noticing about the linkage disequilibrium is that it exists even at $R = .50$. This reflects the second characteristics of optimum models, that even when loci are additive on the primary phenotypic scale, very large amounts of epistasis are generated on the fitness scale because fitness is not monotonic with gene dose. The large epistasis results in linkage effects even with free recombination.

TABLE 3

Stable equilibria for Model I, with different amounts of recombination

R	g_{11}	g_{10}	g_{01}	g_{00}	$p=r$	D'	\bar{W}	\bar{P}	V
.00	.19913	.40043	.40043	.00001	.59956	-1.0000	.95138	5.84670	17.5465
.01	.30506	.34188	.34188	.01118	.64694	-.91038	.94814	6.81564	17.9847
.02	.36381	.30850	.30850	.01919	.67231	-.82129	.94546	7.30789	17.8294
.03	.40313	.28600	.28600	.02487	.68913	-.74265	.94345	7.62661	17.5641
.05	.45275	.25738	.25738	.03250	.71013	-.61327	.94065	8.00729	17.1108
.07	.48238	.24025	.24025	.03712	.72263	-.51746	.93885	8.22940	16.7698
.10	.50913	.22471	.22471	.04144	.73384	-.41501	.93710	8.42475	16.4406
.15	.53288	.21090	.21090	.04532	.74378	-.30968	.93557	8.59495	16.0963
.20	.54575	.20340	.20340	.04744	.74915	-.24600	.93467	8.68589	15.9160
.25	.55400	.19867	.19867	.04866	.75267	-.20483	.93416	8.74475	15.7663
.30	.55950	.19547	.19547	.04956	.75497	-.19121	.93376	8.78303	15.6847
.35	.56350	.19314	.19314	.05022	.75664	-.15213	.93349	8.81086	15.6191
.40	.56650	.19139	.19139	.05072	.75789	-.13477	.93328	8.83166	15.5717
.45	.56900	.18997	.18997	.05106	.75897	-.12101	.93310	8.84954	15.5301
.50	.57050	.18898	.18898	.05153	.75948	-.11098	.93295	8.85796	15.5231

The third important and interesting effect of linkage is shown in the last three columns of Table 3. As for heterotic models, the effect of tightening the linkage is to increase the mean fitness, \bar{W} , of the population. This increase is small however, about 2 percent, although in terms of genetic load it represents a decrease of 26 percent. It is the composition of the mean fitness that changes in an interesting way. Expression (4) shows that the loss of fitness is due to two components: the squared deviation of the population mean phenotype from the optimum, $(\bar{S}-\bar{O})^2$; and the deviations of individuals from the population mean represented by the phenotypic variance, V . As the last two columns of Table 3 show linkage has two opposite effects of fitness. The tighter the linkage, the closer the population mean to the optimum, \bar{O} . This has the effect of making fitness higher for close linkage. However, the tighter the linkage, the greater the within population variance (with the exception of $R = 0$) and this has the effect of lowering the population fitness. The net effect of the two is a slight increase in fitness with increasing tightness of the linkage.

Table 4, giving the results for Model II shows results that are in all respects parallel to those of Model I except that the differences are very much smaller. D' , however, is about as large in Model II as in Model I although the absolute linkage disequilibrium is less. This shows that even when selective differences are small, optimum models generate epistatic deviations that are large in comparison to additive deviations so that relatively severe linkage disequilibria result.

The third quadratic optimum model to be considered is a 5-locus model in which the assumptions are the same as for the 2-locus models. We assume five identical loci determining phenotype additively between loci. The additive effect, $a = 6$, the dominance $h = .8$ and the optimum, $\bar{O} = 24$. This satisfies the requirements for stable equilibria given in Figure 1, since the optimum falls in the shaded region. It is, as a matter of fact, the extension of the two locus case with $a = 6$, $h = .8$ and $\bar{O} = 6$ discussed earlier. that is, $\bar{O} = (n-1)$. Table 5 shows the fitnesses of the different genotypes based upon these parameters. Because all loci are identical only the number of loci homozygous 00 , heterozygous 01 and

TABLE 4
Stable equilibria for Model II with different amounts of recombination

R	g_{11}	g_{10}	g_{01}	g_{00}	$p=r$	D'	\bar{W}	\bar{P}	V
.00	.93382	.03310	.03310	.00000	.96692	-1.00000	.99285	11.6667	.9521
.01	.93642	.03166	.03166	.00026	.96808	-.70130	.99281	11.6789	.9313
.02	.93778	.03090	.03090	.00042	.96868	-.57088	.99279	11.6853	.9198
.03	.93863	.03043	.03043	.00051	.96906	-.47008	.99277	11.6892	.9166
.05	.93965	.02987	.02987	.00061	.96952	-.34444	.99277	11.6940	.9004
.10	.94065	.02932	.02932	.00071	.96997	-.20697	.99277	11.6988	.8841
.20	.94138	.02892	.02892	.00078	.97030	-.11337	.99225	11.7022	.8825
.30	.94163	.02878	.02878	.00081	.97041	-.06853	.99225	11.7033	.8788
.40	.94180	.02869	.02869	.00082	.97049	-.05742	.99225	11.7042	.8757
.50	.94185	.02866	.02866	.00083	.97051	-.04599	.99225	11.7044	.8750

TABLE 5

Fitnesses for Model III. All loci are identical and for each $a = 6$, $h = .8$ and $\bar{O} = 24$

Number of loci with genotype			Fitness
1/1	1/0	0/0	
0	0	5	.98800
0	1	4	.99232
0	2	3	.99568
0	3	2	.99808
0	4	1	.99952
0	5	0	1.00000
1	0	4	.98800
1	1	3	.98272
1	2	2	.97648
1	3	1	.96928
1	4	0	.96112
2	0	3	.89200
2	1	2	.87712
2	2	1	.86128
2	3	0	.84448
3	0	2	.70000
3	1	1	.67552
3	2	0	.65008
4	0	1	.41200
4	1	0	.37792
5	0	0	.02800

homozygous 00 , but not the identity of the loci, need be considered. A homozygote 00 then has a phenotype of 6, a heterozygote 01 a score of 4.8 and a homozygote 11 a score of -6 .

Table 6 is a sample of the results of this model showing equilibrium gametic frequencies, gene frequencies, linkage disequilibrium parameters, variances, and mean fitnesses for selected recombination values. The gametes are given in their binary form, 00000 being a gamete with all five loci represented by the allele with a positive effect on the phenotype and 11111 stands for a gamete with all five loci represented by the allele with a negative effect on the phenotype. Only the gametes with reasonably high frequencies are given in Table 7 and it will be noticed that these are the 16 gametic types with a majority of 0 alleles. Gametes with three or more 1 alleles never exceed a frequency of .00172, at the loosest linkage shown and they all decrease to zero at the tightest linkage. The linkage disequilibrium parameters in the table represent all possible situations, given the fact that the chromosome is left-right symmetrical. Thus, $D'_{12} = D'_{45}$ and $D'_{13} = D'_{35}$, etc. The same symmetry applies to gene $q_1 = q_5$, $q_2 = q_3$. The last column gives the results to be expected if all the loci were at linkage equilibrium, although even a recombination fraction of .50 departs slightly from such an equilibrium situation.

As Table 6 shows, there is very little effect of linkage for recombination greater than .05. As usual, fitness increases with tighter and tighter linkage, but the

TABLE 6
Results of Model III. Symbols are as explained in the text

Gametes	R between adjacent loci											L.E.
	.0000	.0005	.00125	.005	.010	.020	.030	.050	.100	.234	.49040	
00000	.00000	.16908	.28455	.41386	.44838	.46893	.47648	.48284	.48746	.49040	.49300	
00001	.20000	.12430	.09565	.07965	.07717	.07598	.07560	.07531	.07525	.07518	.07491	
00010	.20000	.16069	.13006	.09501	.08617	.08098	.07908	.07748	.07624	.07547	.07491	
00100	.20000	.17457	.14500	.10084	.08904	.08236	.07998	.07800	.07647	.07552	.07491	
01000	.20000	.16069	.13006	.09501	.08617	.08098	.07908	.07748	.07624	.07547	.07491	
10000	.20000	.12430	.09565	.07965	.07717	.07598	.07560	.07531	.07525	.07518	.07491	
00011	.00000	.00509	.00754	.00982	.01048	.01089	.01104	.01117	.01130	.01137	.01138	
00101	.00000	.00901	.01234	.01309	.01254	.01206	.01186	.01168	.01154	.01146	.01138	
01001	.00000	.01154	.01437	.01377	.01290	.01225	.01199	.01176	.01157	.01146	.01138	
10001	.00000	.01225	.01339	.01242	.01197	.01169	.01159	.01151	.01146	.01142	.01138	
00110	.00000	.00573	.00925	.01183	.01189	.01175	.01166	.01156	.01148	.01142	.01138	
01010	.00000	.00966	.01378	.01446	.01347	.01265	.01226	.01193	.01166	.01149	.01138	
10010	.00000	.01154	.01437	.01377	.01290	.01225	.01199	.01176	.01157	.01146	.01138	
01100	.00000	.00573	.00925	.01183	.01189	.01175	.01166	.01156	.01148	.01142	.01138	
10100	.00000	.00901	.01234	.01309	.01254	.01206	.01186	.01168	.01154	.01146	.01138	
11000	.00000	.00509	.00754	.00982	.01048	.01089	.01104	.01117	.01130	.01137	.01138	
D' 12	-1.00000	-.82347	-.65631	-.35265	-.22264	-.12902	-.09050	-.05716	-.03028	-.01312	.00000	
D' 13	-1.00000	-.71442	-.50273	-.21191	-.12326	-.06758	-.04712	-.02902	-.01642	-.00869	.00000	
D' 14	-1.00000	-.61737	-.38674	-.14714	-.08388	-.04608	-.03145	-.01961	-.01246	-.00799	.00000	
D' 15	-1.00000	-.51424	-.28946	-.11085	-.06324	-.03443	-.02385	-.01484	-.01028	-.00686	.00000	
D' 23	-1.00000	-.84587	-.69170	-.36246	-.22535	-.12932	-.09106	-.05699	-.02972	-.01309	.00000	
D' 24	-1.00000	-.73024	-.52027	-.21451	-.12372	-.06769	-.04720	-.02917	-.01629	-.00869	.00000	
$q_1 = q_5$.80000	.83666	.85353	.86363	.86574	.86689	.86729	.86762	.86769	.86777	.86810	
$q_2 = q_4$.80000	.80633	.82222	.84794	.85615	.86142	.86345	.86521	.86659	.86746	.86810	
q_3	.80000	.79502	.80906	.84215	.85316	.85994	.86248	.86463	.86333	.86740	.86810	
\bar{W}	.99414	.99309	.99247	.99180	.99162	.99151	.99148	.99144	.99141	.99140	.99139	
\bar{P}	25.68000	26.15863	26.61655	27.19699	27.36226	27.46270	27.49980	27.53147	27.55430	27.56908	27.58211	
V	14.74427	16.07365	15.75367	14.37925	13.82854	13.47639	13.31799	13.19539	13.13361	13.06167	13.00604	

TABLE 7

Fitnesses for Models IV, V and VI. See text for methods of determining these fitnesses

Number of loci with genotype			Fitnesses		
<i>1/1</i>	<i>0/1</i>	<i>0/0</i>	Model IV	Model V	Model VI
0	0	5	0.0	.0679	.1754
0	1	4	0.0	.1105	.2670
0	2	3	0.0	.1612	.3693
0	3	2	0.0	.2114	.4654
0	4	1	1.0	.2487	.5342
0	5	0	1.0	.2624	.5594
1	0	4	0.0	.1612	.3693
1	1	3	0.0	.2114	.4654
1	2	2	1.0	.2487	.5342
1	3	1	1.0	.2624	.5594
1	4	0	1.0	.2487	.5342
2	0	3	1.0	.2487	.5342
2	1	2	1.0	.2624	.5594
2	2	1	1.0	.2487	.5342
2	3	0	0.0	.2114	.4654
3	0	2	1.0	.2487	.5342
3	1	1	0.0	.2114	.4654
3	2	0	0.0	.1612	.3693
4	0	1	0.0	.1612	.3693
4	1	0	0.0	.1105	.2670
5	0	0	0.0	.0679	.1754
	<i>K</i>		± 1.0	± 1.0	± 2.3
	<i>H</i> ²		1.0	0.2	0.2

effect is extremely small since the population at equilibrium is very close to perfect fitness for all recombination values. As in the previous cases, tight linkage brings the population mean closer to the optimum which accounts in part for the increase in fitness. The changes in gametic and gene frequencies are also reflected strongly in the changes in phenotypic variance. As for the 2-locus models, the phenotypic variance increases steadily with tightening of linkage except for complete linkage where all extreme gametic types are completely absent with a concomitant loss of phenotypic variance. The general increase in phenotypic variance with tighter linkage is in part due to the changing of gene frequencies toward more intermediate values.

The general picture that emerges from treating 5-locus cases such as Model III is that all effects of linkage are the same as in cases of fewer loci but the magnitude of the effects is smaller. This dilution of effects with increasing numbers of loci in the quadratic optimum model is caused by the much higher equilibrium fitness possible with many loci. In addition the gene frequencies at each locus are much closer to fixation at equilibrium and the genetic variance per locus is much smaller as shown in Table 1. Thus, the amount of phenotypic variance that can be maintained in a population is strictly limited although tight linkage can in-

crease this variance. Thus, tight linkage increases the equilibrium phenotypic variance for the 2-locus case by a maximum of 16 percent over the free recombination case, while for the 5-locus model this increase is 24 percent. This variance and the intermediate gene frequencies are maintained at a relatively low level of genetic load.

Double Truncation Models

In the first part of this paper I have shown that truly stable equilibria are difficult to maintain with large numbers of loci with a quadratic deviations model. Moreover, the amount of genetic variance maintained at these equilibria is not great. ROBERTSON (1956) has shown that some other types of optimum model are even worse in this respect in that they predict no stable intermediate equilibria. While linkage does increase the variance maintained at equilibrium in the quadratic deviations model, it does not change the situation materially. It would seem, then, that selection for an intermediate optimum is not likely to account for very much genetic variation in populations. This is not the case however. In this section I will examine the *rate of approach to fixation of genes* for optimum models that do not lead to true stable equilibria. As it will turn out linkage may cause the retention of large amounts of potential genetic variability for extremely long periods although not forever.

The models used to investigate this problem are of the following nature. Five loci determine the genetic score, the loci being all identical in effect, the effects at the various loci adding to each other (no epistasis). At each locus the three genotypes 11, 01 and 00 have the effects a , ah and $-a$ on the phenotypic score. The actual phenotype corresponding to any genotype is assumed to be normally distributed with a mean equal to the genetic score and a variance arising from environmental differences and segregation of genes at other loci than those under investigation. Selection operates by rejecting completely all individuals falling outside the limits $\pm Ka$ and allowing as parents of the next generation all individuals inside these limits. Because each genotype has a normal distribution of phenotypes around its genotypic mean, it is easy with a table of the normal distribution to find the probability that an individual of a given genotype will fall inside the limits $\pm Ka$. This probability is, by definition, the *fitness* or adaptive value of the genotype. If all genotypes have the same environmental variance then clearly the closer the genotypic mean is to zero, the center of the acceptance region, the greater the fitness. When genotypes have different variance, however, the situation may be more complex. These relationships are shown in Figure 2.

The first three double truncation models to be considered are those given in Table 7 as Models, IV, V and VI.

For each model the table gives the truncation limits, K , the initial heritability, H^2 , and the resultant fitnesses of each genotype. In all three models the population is started in linkage equilibrium with gene frequencies .55, .60, .65, .70 and .75 for the five loci in order on the chromosome. This gives an initial genetic variance $\sigma_g^2 = 2.225$. The environmental variance around each genotypic mean is then calculated from the relation $\sigma_e^2 = \frac{2.225 (1-H^2)}{H^2}$

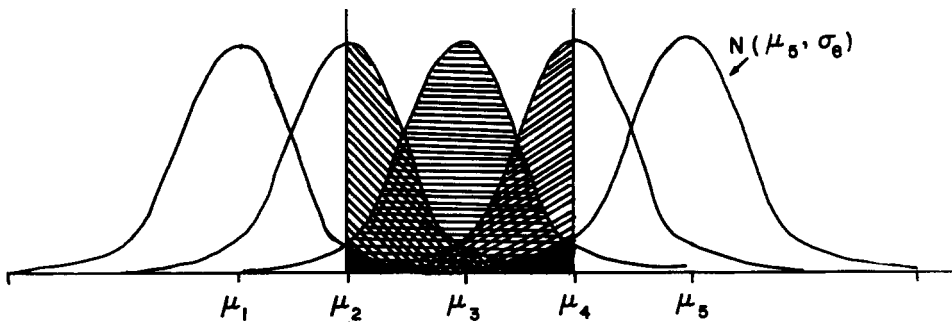


FIGURE 2.—Diagrammatic illustration of double truncation selection. Each normal distribution shown has a mean, μ_i , given by the genotype and a variance caused by environmental variations. All individuals falling between the heavy vertical lines survive while those outside this region do not. The shaded area is the fitness of a genotype.

For Models IV, V and VI H^2 is assumed the same for all genotypes. In Model IV it is 1.0 so that each genotype is either lethal, if it falls outside the limits $\pm K$, or perfectly fit if it falls within these limits. Model V has the same selection limits but $H^2 = .20$ so that 80 percent of the variation is environmental. Model VI decreases the stringency of selection by widening the limits to $K = 2.3$, while holding heritability at 20 percent.

The changes in population structure during the course of selection are shown in Figures 3, 4 and 5. Part A of each figure shows the changes in gene frequencies with time. Part B shows how mean fitness changes and Part C how linkage disequilibrium behaves. To avoid confusion only cases of tight ($R = .01$) intermediate ($R = .05$) and loose ($R = .235$) linkage are given for Parts A and C. As usual these are the linkage values between adjacent genes so that the linkage between outside markers is considerably weaker.

The results shown in Figure 3 are typical of the three cases but the results are more drastic because of the drastic selection differential. At first there is a sharp change in gene frequencies at all loci to bring the mean phenotype of the population close to zero, the center of the acceptable range of phenotypes. This, in turn accounts for the rapid rise in fitness shown in Figure 3B. This process occupies only about two generations and there is no differentiation between cases of loose or tight linkage. After this point, however, the fate of loosely linked and tightly linked genes is very different. Loosely linked genes go to fixation in a symmetrical fashion with one locus temporarily stalled near $q = .50$. After about 90 generations only this locus is still segregating to any marked degree. As the other four loci get arbitrarily close to fixation, the fifth locus will have virtually no selection pressure on it since the heterozygote and two homozygotes at this locus have a fitness of unity. Accompanying this fixation of genes is a slow but steady rise in fitness until it is nearly unity in generation 90. Tightly linked genes are in marked contrast with this picture. Tightly linked loci show a much faster rise in mean fitness so that by generation 15 fitness is in excess of 99 percent. However this rapid increase in fitness is *not* due to gene frequency changes since the

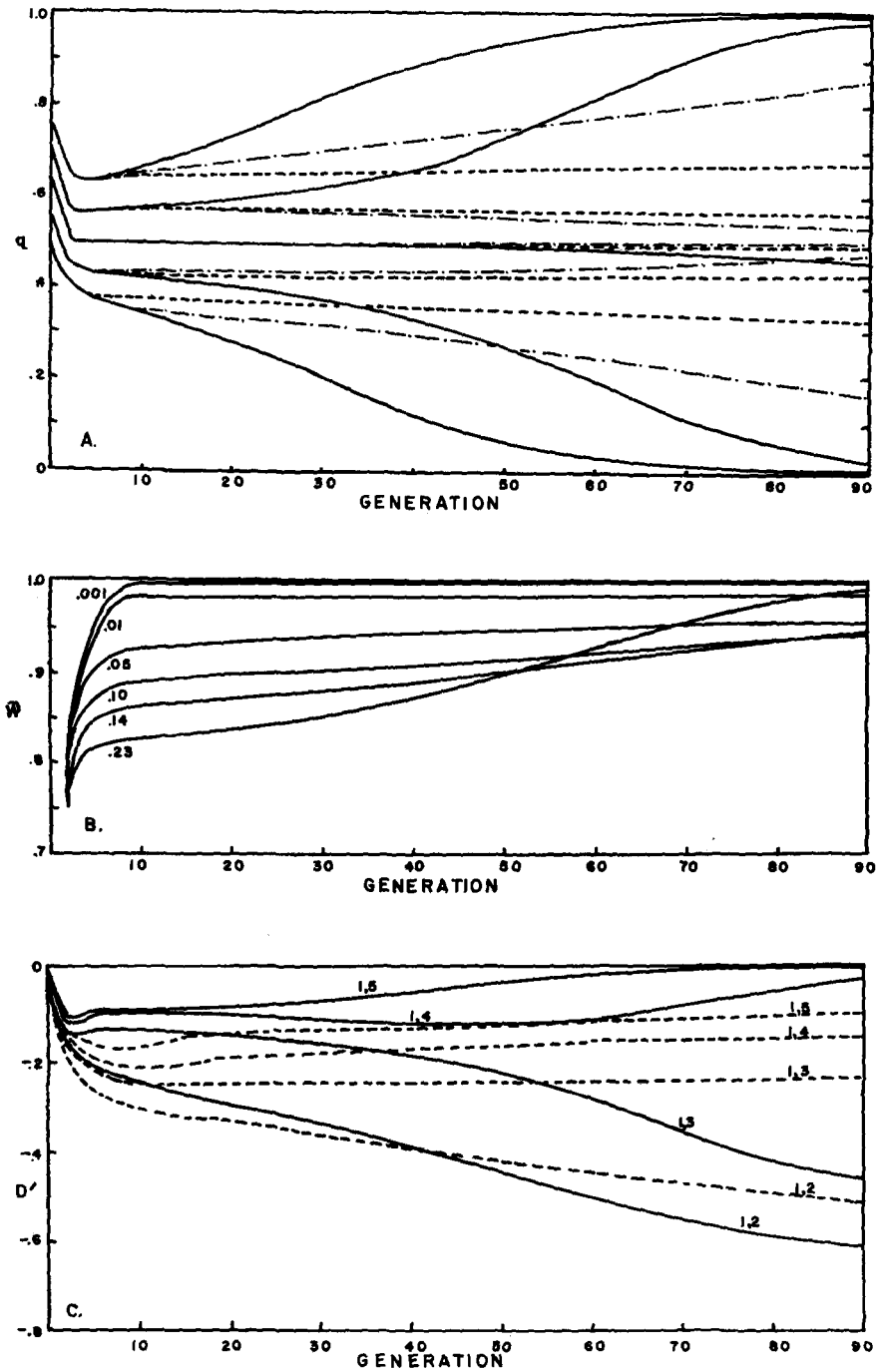


FIGURE 3.—Results for Model IV. Part A shows the frequency of the *O* allele (ordinate) at various times (abscissa) for five loci. Dashed lines are for $R = .01$, mixed line for $R = .05$, solid lines for $R = .234$. Part B shows the mean fitness of the population (ordinate) at various times (abscissa) for different degrees of linkage. Part C shows linkage disequilibrium parameters, D' , (ordinate) at various times (abscissa). Dashed lines are for $R = .01$, and solid lines for $R = .234$.

gene frequencies after becoming symmetrical around .50 remain essentially unchanged for 90 generations. The most deviant gene frequencies at 90 generations are $q_1 = .675$ and $q_5 = .336$, for $R = .01$. With $R = .05$ the three intermediate gene frequencies, q_2 , q_3 and q_4 actually *converge* toward .50 as generations pass, although this is temporary. The rapid increase in fitness has occurred through the fixation of gametic types so that there is a marked deficiency of coupling gametes. This is seen in Figure 3C showing all coefficients of linkage disequilibrium as negative.

One very interesting feature of Figure 3B is the change in the order of mean fitnesses with time. In early generations the tighter the linkage the higher the mean fitness. However, as time goes on the more loosely linked situations slowly increase in fitness passing the more tightly linked case above them. By generation 90 the most loosely linked case is more fit than all except the most tightly linked case, and the next most loosely linked case is beginning its period of more rapid rise. This turnover process eventually results in a complete reversal of the order of fitnesses although this is of very little significance since all cases are very near their maximum possible fitness at that point. As for the quadratic deviations models, loss of mean fitness of the population arises from two causes: deviation of the population mean from the center of the acceptable zone, and phenotypic variance within the population causing individuals to fall outside the acceptance zone. In the case of Model IV all this variance is genetic since the heritability is 1.0, but for Models V and VI some of the variance is nongenetic. Figure 3B can then be used as an index to the genetic variance in later generations. In the first two generations the rapid increase in fitness is due to reducing the deviation of the population mean from the middle of the acceptance range and this is the same for all degrees of linkage. Following that increase, the changes in population fitness are the result of the loss of genetic variance, the highest fitness corresponding to the lowest genetic variance. We see that in the tightly linked case there is extremely rapid loss of genetic variance by the process of *elimination of coupling gametes*; while for loose linkage the loss of variance is slower and occurs by *fixation of genes*. Thus, the relative amounts of genetic variance undergo the same changes in order as do the fitnesses in Figure 3B. Despite the very low genetic variance in the tightly linked cases, the opportunity for the manifestation of new genotypes is much greater because gene frequencies are held at intermediate frequencies. The tightly linked genes then have a greater potential to respond to new selective forces because *potential* genetic variability is maintained in the form of linked complexes.

When we turn to the partial heritability Models V and VI similar features are shown with one exception. In these two models, the partial heritability leads to heterosis of one locus when the other four are fixed. Thus there will be a permanent equilibrium of locus 3 at $q = .50$. This is not of general importance however and arises only because an odd number of loci is involved in a character whose optimum is at 0. The segregation of this locus does not reduce the mean fitness much.

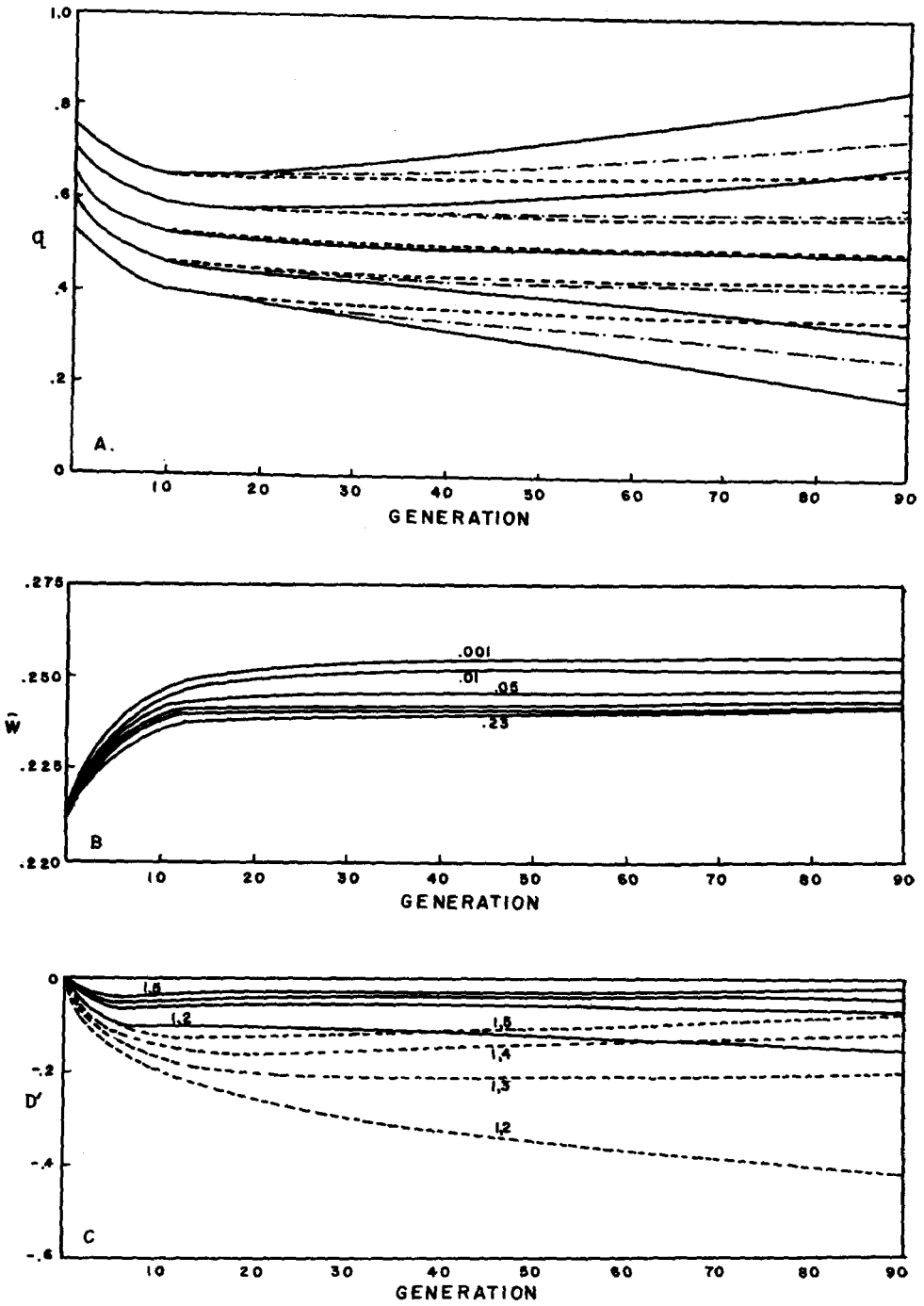


FIGURE 4.—Same as Figure 3 for Model V.

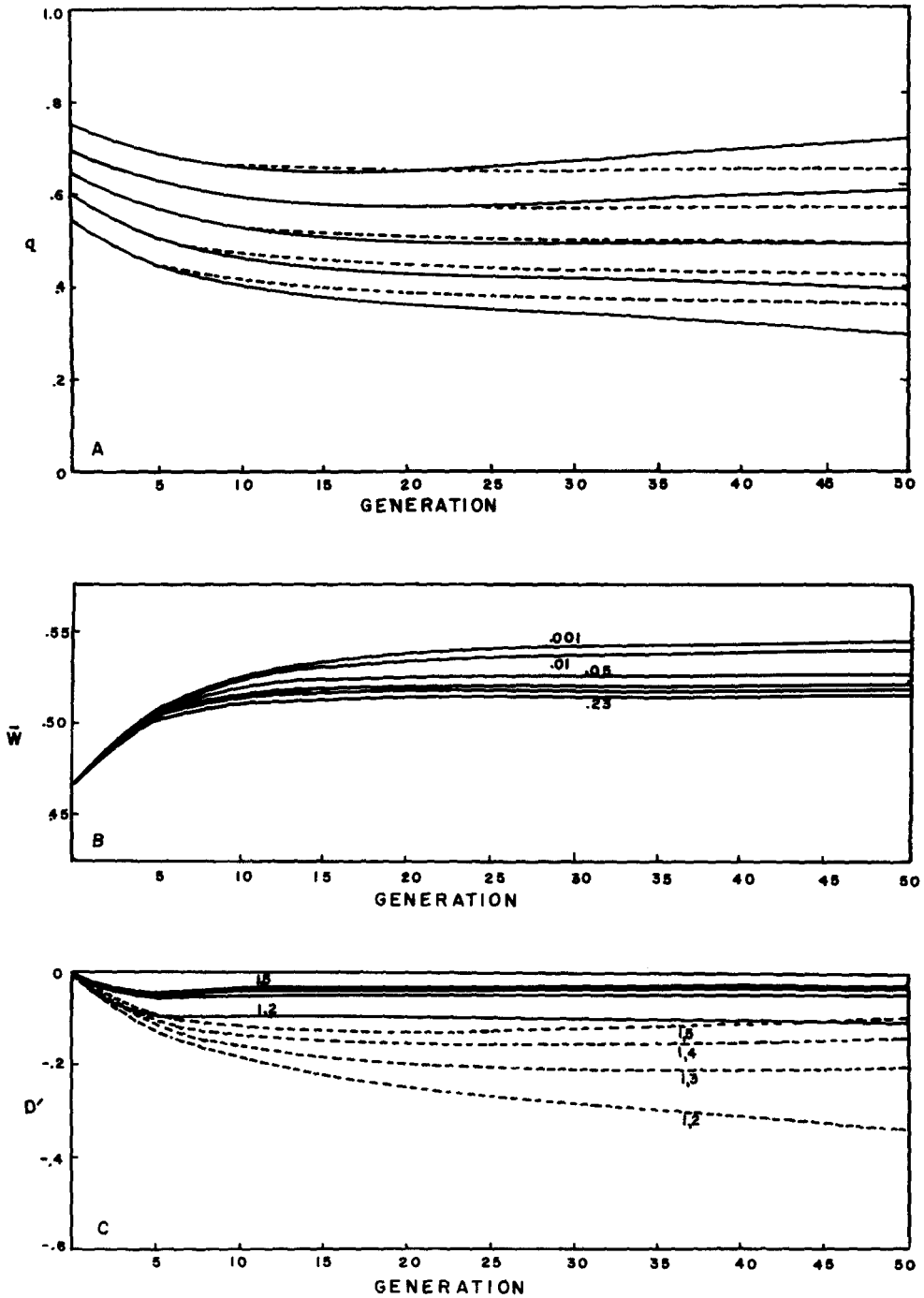


FIGURE 5.—Same as Figure 3 for Model VI.

Homeostatic Models

The last models to be discussed are the same as Models IV, V and VI except that the assumption of equal environmental variance no longer holds. Models VII and VIII whose fitnesses are given in Table 9, are based on LERNER's hypothesis that the greater the degree of heterozygosity, the smaller the environmental sensitivity of a metric trait. In particular Model VII is based upon setting the environmental variances of the genotypes progressively greater for lower

TABLE 8

Environmental variances (σ_e^2) and initial heritabilities in relation to initial genetic variances (σ_g^2) for different genotypes in Models VII and VIII

Number of loci heterozygous	Model VII		Model VIII	
	H^2	σ_e^2/σ_g^2	H^2	σ_e^2/σ_g^2
5	1.000	0.0	1.000	0.0
4	.833	0.2	.500	1.0
3	.667	0.5	.333	2.0
2	.500	1.0	.250	3.0
1	.333	2.0	.200	4.0
0	.167	5.0	.167	5.0

TABLE 9

Fitnesses for Models VII and VIII. See text for method of determining fitness

Number of loci with genotype			Fitnesses	
1/1	1/0	0/0	Model VII	Model VIII
0	0	5	.0789	.0789
0	1	4	.0686	.1103
0	2	3	.0861	.1585
0	3	2	.1685	.2402
0	4	1	.4986	.4102
0	5	0	1.0000	1.0000
1	0	4	.1595	.1595
1	1	3	.2402	.2113
1	2	2	.4102	.2809
1	3	1	.3294	.3646
1	4	0	.4986	.4102
2	0	3	.2251	.2251
2	1	2	.3646	.2634
2	2	1	.4102	.2809
2	3	0	.1685	.2402
3	0	2	.2251	.2251
3	1	1	.2402	.2113
3	2	0	.0861	.1585
4	0	1	.1595	.1595
4	1	0	.0686	.1103
5	0	0	.0789	.0789

degrees of heterozygosity. The initial heritabilities and environmental variances for the two models are given in Table 8. The result of the changing heritability is that the pentuple heterozygote is more fit than any other genotype (Table 9) and this, in turn, results in a stable equilibrium of gene frequencies at $p = q = .50$ for all loci.

Tables 10 and 11 give the stable equilibrium results for various values of recombination. Only 16 gametic types are shown, but their 16 reciprocal types have identical frequencies. What is particularly noteworthy in these results is the appearance of alternating *coupling* and repulsion linkages for tighter linkages ($R = .065-.070$). All other optimum models described in this paper gave only repulsion linkages. The coupling linkages which appear in Models VII and VIII represent the stable configuration for these models and cause a maximization of fitness. Solutions also exist when all linkage are repulsion type but these are *unstable* and result in fitness minima.

I have not given the picture of the early stages of selection for Models VII and VIII since they give stable equilibrium, but these early stages have been computed and a single example is given in Figure 6. Starting in linkage equilibrium, the populations at first produce all repulsion linkages, but as the gene frequencies

TABLE 10

Results of Model VII. Symbols are as explained in the text

Gametes	<i>R</i> between adjacent loci							
	.045	.050	.060	.07	.08	.10	.14	.234
00000	.00016	.00034	.00105	.00179	.00249	.00388	.00666	.01085
00001	.00205	.00351	.00829	.01144	.01328	.01562	.01930	.02351
00010	.00861	.01091	.01474	.01682	.01840	.02000	.02324	.02590
00011	.00328	.00556	.01144	.01525	.01795	.02322	.02765	.03123
00100	.00999	.01245	.01413	.01537	.01716	.01948	.02346	.02623
00101	.06109	.06138	.05046	.04509	.04352	.04261	.04031	.03796
00110	.01276	.01804	.02967	.03440	.03615	.03749	.03725	.03550
00111	.00328	.00556	.01144	.01525	.01795	.02322	.02765	.03123
01000	.00861	.01091	.01474	.01686	.01840	.02000	.02324	.02590
01001	.04706	.05164	.05890	.05774	.05481	.04959	.04380	.03904
01010	.20724	.16610	.10172	.07948	.07041	.05943	.04933	.04147
01011	.01109	.06138	.05046	.04509	.04352	.04261	.04031	.03796
01100	.01276	.01804	.02967	.03440	.03615	.03749	.03725	.03550
01101	.04706	.05164	.05890	.05774	.05481	.04959	.04380	.03904
01110	.01290	.01905	.03612	.04188	.04198	.04016	.03745	.03516
01111	.00205	.00351	.00829	.01144	.01328	.01562	.01930	.02351
<i>D'</i> 12	-.59512	-.52904	-.43516	-.37836	-.33236	-.25792	-.17792	-.11036
<i>D'</i> 13	+.35244	+.24136	+.14528	-.02228	-.04400	-.06264	-.06588	-.05652
<i>D'</i> 14	-.24484	-.16040	-.05548	-.03840	-.03752	-.04700	-.04876	-.04784
<i>D'</i> 15	+.09212	+.02380	-.03264	-.03616	-.03652	-.04832	-.04844	-.05400
<i>D'</i> 23	-.64448	-.54976	-.32604	-.23696	-.20656	-.17768	-.14140	-.10120
<i>D'</i> 24	+.42628	+.31088	+.08208	+.00636	-.01848	-.04236	-.05548	-.05332
\bar{W}	.43673	.42021	.39785	.38979	.38498	.37894	.36995	.36099

TABLE 11

Results of Model VIII. Symbols are as explained in the text

Gametes	<i>R</i> between adjacent loci							
	.005	.01	.065	.07	.08	.10	.14	.234
00000	.00000	.00000	.00208	.00334	.00508	.00767	.01206	.01713
00001	.00000	.00000	.00892	.01241	.01555	.01865	.02261	.02628
00010	.00003	.00012	.01686	.01909	.02133	.02322	.02618	.02809
00011	.00000	.00000	.01159	.01555	.01960	.02504	.02883	.03125
00100	.00003	.00012	.01585	.01743	.02003	.02295	.02655	.02838
00101	.00457	.00921	.04915	.04479	.04222	.04060	.03794	.03565
00110	.00003	.00014	.02778	.03253	.03523	.03642	.03565	.03402
00111	.00000	.00000	.01159	.01555	.01960	.02504	.02883	.03125
01000	.00003	.00012	.01686	.01909	.02133	.02322	.02618	.02809
01001	.00436	.00879	.05649	.05527	.05112	.04535	.03993	.03622
01010	.48197	.46321	.10992	.08301	.06627	.05372	.04400	.03782
01011	.00457	.00921	.04915	.04479	.04222	.04060	.03794	.03565
01100	.00003	.00014	.02278	.03253	.03523	.03642	.03565	.03402
01101	.00436	.00879	.05649	.05527	.05112	.04535	.03993	.03622
01110	.00003	.00013	.03057	.03693	.03850	.03713	.03512	.03365
01111	.00000	.00000	.00892	.01241	.01555	.01865	.02261	.02628
<i>D'</i> 12	-.98136	-.96160	-.42476	-.35720	-.28544	-.20172	-.12544	-.07180
<i>D'</i> 13	+.96380	+.92580	+.08748	+.01024	-.02966	-.05016	-.04912	-.03788
<i>D'</i> 14	-.94652	-.89128	-.06548	-.03944	-.03324	-.03928	-.03660	-.03204
<i>D'</i> 15	+.92860	+.85592	-.00924	-.02420	-.02800	-.03712	-.13448	-.03520
<i>D'</i> 23	-.98220	-.96324	-.34716	-.24988	-.19212	-.15152	-.10804	-.06832
<i>D'</i> 24	+.96468	+.92756	+.09828	+.02044	-.01828	-.04020	-.04464	-.03672
\overline{W}	.59257	.57275	.34583	.33754	.33158	.32637	.32058	.31612

reach equilibrium, some of these repulsion linkages disappear and are replaced by coupling phase linkages. The particular linkages shown in Table 10 are not the only possible equilibria. There are ten alternative configurations in which the majority of gametic types form a pair of the type 11100/00011 or 10110/01001, or, etc. All these equilibria have the same mean fitness. As Table 10 shows there are very considerable linkage disequilibria between adjacent genes even when these recombine as much as 6 percent. When adjacent recombination fractions are 1 percent, the outside markers, which are then 4 map units apart, are in very great disequilibrium ($D' = .84$ for Model VIII).

The most remarkable feature of these models is the very large increase in fitness that results from linkage. In Model 8, if the adjacent genes recombine 1 percent the average equilibrium fitness is .57275 compared with fitness of .31612 when $R = .24$. This is an increase of 81 percent in fitness and if the genes are really tightly linked ($R = .001$) fitness is essentially doubled.

General Implications of the Results

As D'ARCY THOMPSON and HALDANE long ago observed, there is a strong pressure of natural selection toward intermediate values of various metric traits

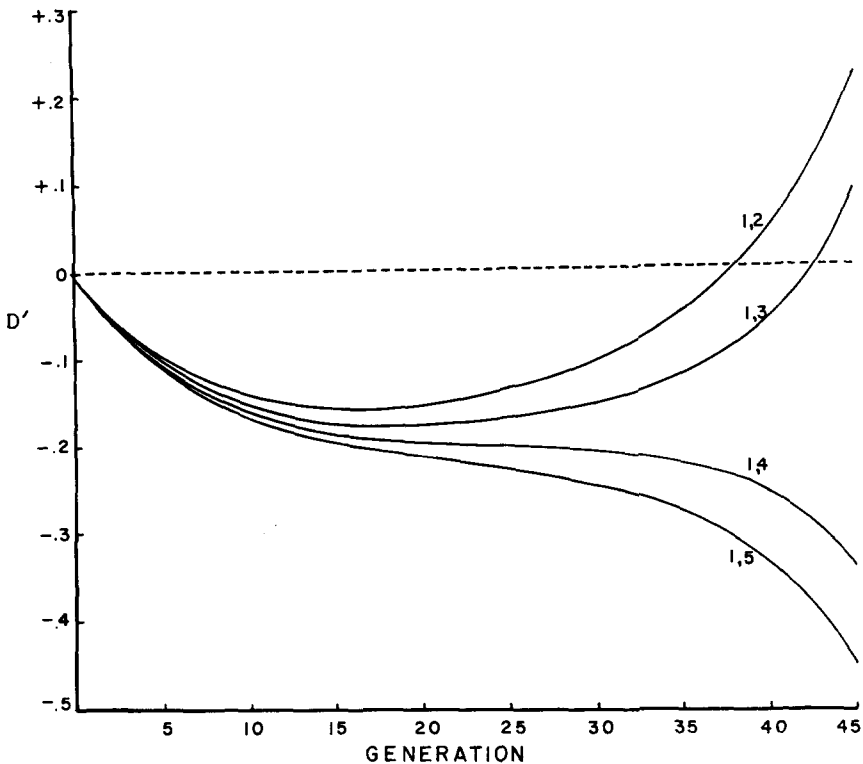


FIGURE 6.—Linkage disequilibrium parameters, D' , (ordinate) at various times (abscissa) for Model VIII with $R = .001$.

of organisms. In fact, with exception of components of fitness such as make up the age specific mortality and fecundity schedules, it is difficult to imagine any character of a plant or an animal in which the most extreme size is of advantage. For this reason the consequences of various schemes of optimizing selection are of great interest in population genetics. The question is to what extent such optimizing selection results in the long-time maintenance of genetic variability.

When genes are freely recombining, optimizing selection does not usually lead to stable genetic variation. Of the models examined in this paper only those in which multiple heterozygotes are better buffered against environment lead to much stable variation. The linkage of genes, however, leads to rather a different result. The standing genetic variance in an equilibrium population is made larger by linkage in quadratic models but smaller in double truncation models, but the possibility of long term maintenance of gene frequencies at intermediate values is greatly enhanced. Linkage of genes may result in the much more rapid *immediate* response to selection with an accompanying exhaustion of genetic variance, while at the same time maintaining the possibility of *future* genetic variation by preventing the fixation of gene frequencies. For environments that are stable over long periods but undergo occasional radical shifts this is an opti-

mal genetic strategy. On the other hand, in environments undergoing rapid frequent fluctuation, a less labile system is optimal and tight linkage is at a disadvantage.

The other point of general interest is the degree to which populations may be permanently out of linkage equilibrium. In the earlier papers on this subject KOJIMA and I have pointed out that interaction between loci in determining fitness is necessary if any permanent linkage disequilibrium is to result. By interaction we have meant departures from *additivity* between loci and this follows from our particular model of natural selection in discrete generations. When a continuous generation model is used, the definition of interaction is the deviation from *multiplicativeness* between loci, or additivity on a logarithmic fitness scale (KIMURA 1956). Thus, whether or not a particular scheme of fitnesses leads to linkage disequilibrium depends in part on the model of the breeding structure. For optimizing models, however, interaction is present irrespective of the discreteness of generations because the fitnesses do not bear a simple relationship to the rest of the phenotype. On either discrete or continuous generation models, optimizing selection results in very large amounts of interaction on the fitness scale and this, in turn, results in linkage disequilibrium for even the loosest linkages. To the extent that genes are selected to produce optimum intermediate phenotypes, and this must be a common mode of selection, to that extent linkages play an important role in the genetic change in the population and linked complexes of genes, out of equilibrium, should be found in natural populations.

Over the past eight years I have had many occasions to discuss these matters with Dr. KENICHI KOJIMA and I am deeply grateful for the illuminating insights he has given me.

SUMMARY

This paper is concerned with the genetic structure of populations subject to "optimizing" selection, that is, selection favoring some intermediate phenotype and discriminating against phenotypes at both extremes. A number of specific models have been examined in which the fitness of phenotypes falls off with increasing deviation from the optimum phenotype. For freely recombining genes virtually no optimum model leads to a maintenance of significant amounts of genetic variation in a permanently stable state. An exception to this rule is a model in which multiple heterozygotes have a lower environmental sensitivity than other genotypes. When recombination is restricted, however, long-time quasi-stable equilibrium of gene frequencies result. The rate of increase of fitness with time before the quasi-stable equilibrium is reached is also strongly affected by linkage with tight linkage causing rapid initial increases in fitness but lower rates of increase in later generations.

Pronounced linkage disequilibrium also appears during selection and at equilibrium. This is usually in the form of excesses of repulsion linkages, but in some cases there is an alternation of coupling and repulsion linkages along the chromosome similar to the "relational balance" hypothesis of MATHER.

APPENDIX I. *Proof that shaded area in Figure 1 represents the necessary condition for stable equilibrium of n loci.*

Let \bar{O} = optimum. P_j = mean genotypic value due to j th locus. \bar{P} = mean genotypic value. a = additive effect per locus. h = dominance effect per locus. Then KOJIMA (1959) shows that a necessary condition for stability is that for all j

$$(1) \quad \bar{P} - P_j + \frac{a(h-1)}{2} < \bar{O} < \bar{P} - P_j + \frac{a(h+1)}{2}$$

Now (1) specifies the following three properties: (a) \bar{O} lies in an interval whose width is a ; (b) The upper and lower limits of \bar{O} are each linear in h with slope $a/2$; (c) Lower limit of $\bar{O} = \bar{P} - P_j$ when $h = 1$. We thus have a band of the shape and slope shown in Figure 1 for each locus relating the dominance h to the optimum \bar{O} . Since we are looking for a *necessary* condition we want the simultaneous condition on all the j loci which is as broad as possible. But this will be given by the intersection of all the conditions and will be as broad as possible when all bands are identical and include the same region of the h, \bar{O} plane since they are all of fixed width a and fixed slope $h/2$. This will occur when

$$(2) \quad \bar{P} - P_i = \bar{P} - P_j$$

for all i and j .

Since $\bar{P} = \sum_{j=1}^n P_j$, if (2) is true then

$$(3) \quad \bar{P} - P_j = nP_j - P_j = (n-1)P_j .$$

Finally, when $h = 1$, $P_j = a$ because of fixation of the plus allele. Thus, from (3) and condition *c* above

$$(4) \quad \bar{O} = (n-1)a \text{ when } h = 1.$$

Conditions a, b and c in conjunction with (4) are the required area.

APPENDIX II. *Derivation of limiting phenotypic variance maintained by the quadratic optimum model*

Expression 7a of the text states that at equilibrium

$$(1) \quad a(h^2-1)(1-2q_j) + 2[1 + h(1-2q_j)][\bar{P} - P_j - \bar{O}] = 0$$

for all j .

Taking the case where all loci have the same gene frequency, which will maximize the variance, (1) becomes, for n loci

$$(2) \quad a(h^2-1)(1-2q) + 2[1 + h(1-2q)][(n-1)P_j - \bar{O}] = 0$$

Setting $q = 1 - \epsilon$ and ignoring higher powers of ϵ ,

$$(3) \quad P_j = aq^2 + 2pqah - p^2a = a[1 - 2\epsilon(1-h)]$$

so that from (2) and (3)

$$(4) \quad a(h^2-1)(2\epsilon-1) + 2[1 + h(2\epsilon-1)][(n-1)a - \bar{O} - 2\epsilon(1-h)(n-1)a] = 0$$

It was proved in Appendix I that $(n-1)a - \bar{O}$ is a small constant for all values of n . Therefore, as n grows large without bound the left hand side of (4) will also grow without bound unless $n\epsilon$ is bounded.

Therefore

$$(5) \quad \lim_{n \rightarrow \infty} \varepsilon = 0$$

Setting $\bar{O} - (n-1)a = aE$ and rearranging (4) yields

$$(6) \quad \varepsilon = \frac{a(h^2-1) + 2(1-h)aE}{2a(h^2-1) - 4aDh - 4(1-h)^2(n-1)a}$$

$$(7) \quad \lim_{n \rightarrow \infty} n\varepsilon = \frac{(1+h) - 2E}{4(1-h)}$$

Finally, the variance at a single locus is (from text equation 5)

$$(8) \quad \sigma_j^2 = 2a^2(1-h)^2\varepsilon$$

ignoring higher powers of ε . Therefore the limiting variance of total phenotype from (7) and (8) is

$$(9) \quad \sigma^2 = \lim_{n \rightarrow \infty} n\sigma_j^2 = 2a^2(1-h)^2 \lim_{n \rightarrow \infty} n\varepsilon = a^2(1-h) \left(\frac{1+h}{2} - E \right).$$

LITERATURE CITED

- BODMER, W. F., and P. A. PARSONS, 1962 Linkage and recombination in evolution. *Advan. Genet.* **11**: 1-100.
- FISHER, R. A., 1950 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- KIMURA, M., 1956 A model of a genetic system which leads to closer linkage by natural selection. *Evolution* **10**: 278-287.
- KIMURA, M., and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. *Genetics* **49**: 725-738.
- KOJIMA, K., 1959a Role of epistasis and overdominance in stability of equilibria with selection. *Proc. Natl. Acad. Sci. U.S.* **45**: 984-989. — 1959b Stable equilibria for the optimum model. *Proc. Natl. Acad. Sci. U.S.* **45**: 989-993.
- KOJIMA, K., and T. KELLEHER, 1963 A comparison of purebred and crossbred selection schemes with two populations of *Drosophila pseudoobscura*. *Genetics* **48**: 57-72.
- LERNER, I. M., 1954 *Genetic Homeostasis*. Wiley, New York.
- 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., 1941 Variation and selection of polygenic characters. *J. Genet.* **41**: 159-193.
- ROBERTSON, A., 1956 The effect of selection against extreme deviants based on deviation or homozygosis. *J. Genet.* **54**: 236-248.
- WRIGHT, S., 1935 Evolution in populations in approximate equilibrium. *J. Genet.* **30**: 257-266.