GENETIC DIVERSITY AND DIVERSITY OF ENVIRONMENT: MATHEMATICAL ASPECTS

HOWARD LEVENE
COLUMBIA UNIVERSITY

1. Introduction

This paper is more mathematical than that of Professor Dobzhansky [3]. However, it is similar in raising more questions than answers. It is designed to present the same set of problems in a way that it is hoped will be somewhat more precise and "comfortable" to mathematicians, although time and space do not permit any very detailed presentation except for a few elementary examples.

We shall be concerned mainly with natural selection. There are two main kinds, intrapopulational and interpopulational. Most attention has gone to intrapopulational selection, which is highly opportunistic and short sighted. An individual at a certain stage in the life cycle leaves more offspring that reach the same stage in the life cycle, and thus its genes are increased in number in the next generation. If the selective pressures keep changing from generation to generation, the population keeps trying to be adapted to conditions prevailing in the previous generation. Furthermore, there is no way for a gene or gene combination that will be good at some future time, say of catastrophe, but leaving relatively fewer descendants now, to become established. If, however, it can hang on, even at low frequency until the catastrophe, it may become established.

On the other hand, all populations and species we see now are descendants of populations that have survived all of geologic time. The populations that opportunism drove into dead ends are dead, and those that solved short term problems in a way that happened to be good over the long term survived. This is the process of long term interpopulation selection. The mathematics of interpopulation selection has been largely ignored. Selection between species was considered by Voltera, Lotka, and others, and between potentially interbreeding Mendelian population by Wright [38] to some extent. Levins, [16] to [20], and Lewontin [22], [24] have considered optimal strategies for populations. Those populations that hit on good strategies will be preserved, so that existing populations should mirror, in many ways, the optimal strategies. Even if optimal

Research supported by the Office of Naval Research under Contract Nonr-4259(08), Project NR-042-034 at Columbia University. Reproduction in whole or part is permitted for any purpose of the United States Government.

strategies are rarely attained by intrapopulation selection, interpopulation selection will insure that they are greatly overrepresented in living populations.

We find that many or most Mendelian populations have genetic diversity or polymorphism, so this must be either a good strategy *per se*, or be a byproduct of a good strategy. Professor Dobzhansky has suggested reasons why diversity may be good in the long run, and Levins has studied mathematical conditions under which polymorphism is advantageous. In the present paper we consider mainly the genetic mechanisms that lead to genetic diversity in populations by intrapopulation selection.

2. Intrapopulation selection and mutation

In the earliest work on intrapopulation selection, Fisher [3], Haldane [4], and Wright [29], the mathematical treatment was of one locus with alleles $A_1, \cdots, A_I$ present with gene frequencies $(q_1, \cdots, q_I) = q$ and considered the stochastic process followed by the vector q in time, usually for $I = 2$. This was extended early in principle to J loci, the jth locus having I_j alleles, and the population being specified by the random matrix (q_{ji}), where q_{ji} is the frequency of the ith allele at the jth locus, and $\sum_i q_{ji} = 1$, but very few actual solutions were obtained. It now turns out that this model has limited applicability and that we must specify gametic frequencies $q_{i_1 i_2 \cdots i_J}$ of gametes $A_{1i_1}A_{2i_2}, \cdots, A_{Ji_J}$ so that the number of q is ΠI_j rather than $\sum I_j$. This is because the theoretical relation of equilibrium in the absence of selection

$$q_{i_1 i_2 \cdots i_J} = \prod_j q_{ji_j} \tag{2.1}$$

no longer holds if there are specific interactions in selective value between loci, so that the selective value of genotype $A_{1i_1}A_{1i'_1}, \cdots, A_{Ji_J}A_{Ji'_J}$ cannot be predicted from the selective values of $A_{1i_1}A_{1i'_1}$, $A_{2i_2}A_{2i'_2}$, and so forth, considered individually. This more complicated model is only now being satisfactorily attacked, in part by Lewontin, who participated in this Symposium [25]. Finally, there are cases, largely unexplored, where the population must be specified by the frequency of diploid genotypes, giving $(\Pi I_j)(1 + \Pi I_j)/2$ parameters q.

All these cases can be attacked on various levels. If the population is considered infinitely large, with constant selection pressures, the process becomes deterministic. We can then ask whether there is an equilibrium with not all of the q equal to 0 or 1, which q reach fixation at 0 and at 1, and what is the rate of approach to the final state. Even here gaps are bigger than our knowledge. If the population is finite, all of the q reach fixation with probability unity, unless there is mutation or migration from an infinite population, when there is a nondegenerate stable equilibrium. The expected rate of approach to fixation, in the one case, and the stable distribution in the other, have been found explicitly for certain special cases, and in symbolic form in more general cases. (See, for example, Feller [4], Fisher [5], Kimura [12], Moran [30], Wright [33],

[36], [38], and the papers by Bodmer, Karlin, and McGregor [1], [9] in this Symposium.)

If the population is considered as infinite, with no mutation or migration, and selective values are random variables, there is the phenomenon which Kimura [12] called quasifixation, in which the probability distribution of the gene frequency q becomes concentrated very close to zero and one. This becomes more intuitive if one looks at $z = \log [q(1 - q)]$. Then z tends to $-\infty$ or $+\infty$, as q tends to 0 or 1, and quasifixation is simply escape to infinity. The stochastic process for z is essentially a random walk in the discrete generation case, and a Wiener process in the continuous case. On the case of most interest, where the mean displacement of z is zero, the probability that z will be in any fixed finite interval tends to zero with time, but, nevertheless, the process does return to the origin infinitely often with probability one. Thus, quasifixation differs from true fixation in that it is certain to end ultimately. In both the case of fixation and quasifixation there may be limiting distributions conditioned on $a \leqq q \leqq b$ for fixed $a > 0$ and $b < 1$.

Mathematically, the problem of genetic diversity within a population is whether there is a nontrivial stable equilibrium or stable steady state distribution. In some cases there may be a nontrivial equilibrium to which the population usually returns when disturbed, even though it eventually escapes to fixation or quasifixation. For example, there may be strong selection pressures of the sort to be discussed below which would lead to equilibrium in an infinite population. In a finite population with no mutation ultimate fixation is certain, but selection will usually return q to its equilibrium value after any random departure. Such quasiequilibria are of mathematical interest, but practically are true equilibria and become so with more realistic models.

A very low rate of mutation, of such a nature that it is possible to go from any allele at a locus to any other in a bounded number of steps, is sufficient to make the process of change of gene frequencies ergodic, and insure a stable equilibrium or distribution. In nature, mutation rates are usually small, of the order of 10^{-5} or 10^{-6}. Thus, in a small population the steady state distribution may assign very large probabilities to the "trapping states" 0 and 1, and a population reaching one of these states may stay there a long time before a chance mutation starts the process of fluctuation in motion again. Furthermore, even in infinite populations, selective forces are usually so much bigger than mutational forces, that mutation acting against directed selection can maintain only a very low gene frequency for the allele that is selected against. In higher organisms there are tens of thousands of loci and a large number of these produce deleterious alleles by mutation. While each such allele will have a very low frequency in the population, the sum of all variation so generated at all loci may be considerable. However, in general, mutation is not an interesting source of polymorphism, even though it is the ultimate source of all genetic diversity. Accordingly, we shall turn our attention to various kinds of

selection. Also, because the deterministic case is much better understood than the stochastic case, we shall confine our attention in the sequel to the deterministic case except when otherwise specified.

3. Simple selection

To fix our ideas, let us consider three simple models of natural selection that will recur in our further discussions. The first is genic selection with two alleles, which is equivalent to selection in haploids or to selection between two species. There are two alleles, A_1 and A_2, at a given locus, each individual carrying one of them; the alleles having frequencies q and $1 - q$, respectively. If the expected number of offspring of A_1 and A_2 individuals are W_1 and W_2, the value of q in the next generation is

$$q' = \frac{W_1 q}{W_1 q + W_2(1 - q)}. \tag{3.1}$$

Since this is homogeneous of degree zero in W, any numbers proportional to W_1 and W_2 will do, and it is customary to set one of the W equal to 1, say $W_1 = W$, $W_2 = 1$. These W (or W_1 and W_2) are called adaptive values or selective values. It is often more convenient to let $W = 1 - s$, where s is called a selective coefficient, or $W = e^m$, where m is called a Malthusian parameter. If selection is weak, $W \sim 1$ and s and m are close to zero, furthermore, $-s = m + m^2/2 + \cdots \sim m$. If $q(T)$ is the value of q after T generations, this case permits the simple solution

$$q(T) = \frac{q(0)W^T}{[1 - q(0)] + q(0)W^T}. \tag{3.2}$$

There is no equilibrium except the trivial neutral equilibrium $q(T) = q(0)$ when $W = 1$.

The second model involves a diploid locus with two alleles, one recessive,

TABLE I

RESULTS OF RANDOM MATING UNDER MODEL 2
(Diploid locus with two alleles, one recessive and one dominant)

Genotypes	aa	aA	AA	Sum
Frequency before selection	q^2	$2q(1 - q)$	$(1 - q)^2$	1
Adaptive values	$(1 - s) = W$	1	1	
Numbers proportional to frequency after selection	$(1 - s)q^2$	$2q(1 - q)$	$(1 - q)^2$	$\overline{W} = 1 - sq^2$

one dominant. With random mating we have the results shown in table I. The gene frequency in the next generation is

$$q' = [(1 - s)q^2 + 2q(1 - q)/2](W)^{1/2} = q(1 - sq)/(1 - sq^2). \tag{3.3}$$

There is no simple formula for $q(T)$ except for the case of a recessive lethal where $W = 0$, giving $s = 1$, and

$$q(T) = \frac{q(0)}{1 + Tq(0)}. \tag{3.4}$$

Again, there is no equilibrium except a neutral one when $W = 1$.

The third model is the most general one for a single locus with random mating. There are I alleles, the ith having frequency q_i, and genotype A_iA_j has adaptive value $W_{ij} = W_{ji} = 1 - s_{ij}$. Then, after one generation,

$$q_i' = \frac{\sum_j q_iq_jW_{ij}}{\overline{W}}, \tag{3.5}$$

where

$$\overline{W} = \sum\sum q_iq_jW_{ij} \tag{3.6}$$

is the weighted mean adaptive value. It can be shown that $\overline{W}'$, the mean adaptive value using the q_i', is greater than $\overline{W}$ unless the gene frequencies are stationary, so that equilibrium points may be found by maximizing $\overline{W}$. For general I there can be at most one stable equilibrium point with all $q_i > 0$; its existence involves certain matrix conditions (see, for example, Levene, Dobzhansky, and Pavlovsky [12], and Mandel [24]); there is no simply verifiable condition on the W. However, for $I = 2$ there is an equilibrium with

$$q_1 = \frac{(W_{12} - W_{22})}{(W_{12} - W_{22}) + (W_{12} - W_{11})}, \tag{3.7}$$

which is stable if and only if $W_{12} > W_{11}$ and $W_{12} > W_{22}$; that is, the heterozygote is superior to both homozygotes. This is called the condition of overdominance, or heterosis, or hybrid vigor. For more alleles the W_{ij} must roughly be bigger than the W_{ii} but in precise terms this is neither necessary nor sufficient.

4. Multiple niches

We now consider the behavior of these simple models when there are multiple ecological niches; that is, a number K of subpopulations each with different values of the W. Again there are various models. The present author [13] first considered the case where the kth niche contributes a constant fraction of individuals c_k to the next generation, where after selection individuals from all niches form one big random mating population, and where the newly formed zygotes are assigned to niches at random regardless of genotype. For genic selection

$$q' = \sum_{k=1}^{K} c_kq_k', \tag{4.1}$$

where

$$q'_k = \frac{W_k q}{W_k q + (1 - q)}, \tag{4.2}$$

and the adaptive values in the kth niche are W_k and 1. There will be a stable equilibrium if $\sum c_k W_k > 1$ and $\sum c_k / W_k > 1$, or in other words if $\bar{W} > 1$ and $\tilde{W}_h < 1$, where $\bar{W}$ is the mean of the W_k weighted by the niche sizes c_k but not by the q and $\tilde{W}_h$ is the correspondingly weighted harmonic mean of the W_k. The same result holds for a recessive.

For the diploid case with two alleles and $W_{11} = W$, $W_{12} = 1$, $W_{22} = V$, there will be a stable equilibrium if $\tilde{W}_h < 1$ and $\tilde{V}_h < 1$, a weaker condition than the one we might expect that $\bar{W} < 1$ and $\bar{V} < 1$. Furthermore, the condition on the harmonic means is sufficient but not necessary. No simple conditions are known for more than two alleles.

Dempster [2] has pointed out that the above results do not hold if the c_k depend on the value of q. In particular, if the adaptive values W_{ijk} in the kth niche are considered as absolute expected number of offspring rather than as relative values, conditions for equilibrium are no weaker than in the one niche case. In fact this model is equivalent to a single niche model with $W_{ij} = \sum_k W_{ijk}$.

My model was chosen deliberately to show increased opportunity for equilibrium even in an unfavorable case. There are two simple ways of making it more realistic and also increasing the opportunity for equilibrium. One which seems not to have been treated mathematically is to permit zygotes to select the niche where they will develop. For example, with the genic selection model and two alleles, suppose there are two niches, that haploids of type A_1 have probability α_1 of going to niche 1 and $1 - \alpha_1$ of going to niche 2, while A_2 have probability $1 - \alpha_2$ of going to niche 1 and α_2 for niche 2. If $\alpha_1 = \alpha_2 = 1/2$, we have the previous model with $c_1 = c_2 = 1/2$. If $\alpha_1 = \alpha_2 = 1$, there will be equilibrium as long as A_i can survive in niche i, $i = 1, 2$. Suppose that 1 is the better niche for A_1 and 2 for A_2, and α_1 and $\alpha_2 > 1/2$, there will be increased opportunities for equilibrium. This model can be extended to diploids with two alleles, but we must then decide whether to give heterozygotes their own niche, or divide them up in some way, and also whether to assume a Levene or Dempster type model. One can make guesses as to the results, but no algebraic analysis has yet been done.

Moran [29] considered the case of two populations essentially of equal size where mating is random *within* each population, but after reproduction there is migration so that the next generation in each population consists of a fraction $1 - m$ from that population and a fraction m from the other. If $m = 1/2$, this reduces to Levene's case with $c_k = 1/2$, while if $m = 0$, each population goes its own way. Moran only considered the case where W_{12} is halfway between W_{11} and W_{22}. This is essentially the same as genic selection. He further specialized the model to the symmetric case where W_{11} and W_{22} are interchanged in the two populations. Thus, if there is an equilibrium, the average q in the two

populations is 1/2. Moran finds that there is always a stable equilibrium if $m < 1/2$. Letting $W_{11} = 1 - s$, $W_{12} = 1$, $W_{22} = 1 + s$, the equilibrium value of q will be near 1/2 in each niche if $m \gg s$, while it will be near 0 in one and 1 in the other if $m \ll s$. The algebra for this symmetric case is not trivial, and while Moran says his results can be extended to different strengths of selection in the two populations as long as selection is in different directions in the two, and heterozygotes are intermediate; and to unequal migration in the two directions, he gives no details. Extension to more than two populations introduces completely new problems as to whether all m should be small but equal, or migration should be greater from "nearby" populations (suitably defined) or only from "neighbors" plus perhaps a fraction from the mean of all populations. Such models have been studied by Wright [37], Malécot [27], Pollak [31], Weiss and Kimura [32], and by Moran in this same paper, but from the stochastic point of view, and, for the most part, with selection absent or identical in all subpopulations. Here, particularly, there is still much to do.

The case of one population, but W changing from one generation to the next was considered by Dempster [2] and by Haldane and Jayakar [7]. For genic selection there is no equilibrium except the neutral one when $\Pi W_i = 1$. However, for a recessive there is equilibrium when

$$\sum^{T} W_t/T > 1, \quad (\prod^{T} W_t)^{1/T} < 1, \tag{4.3}$$

that is, the arithmetic mean of the $W_t > 1$ and the geometric mean of the $W_t < 1$.

For the general 2 allele case with $W_{12} = 1$ there will be equilibrium if the geometric means of the W for the homozygotes are both less than 1. Where the temporal change is due to the W_{ijt} being independent observations on the same random variables H_{ij}, the effect of the variance of W is to make equilibrium more difficult to attain than if $W_{ij} \equiv E(H_{ij})$.

Another case of considerable interest is that in which W for a genotype is a decreasing function of the frequency of that genotype. This would occur, for example, in niche selection, considered above, if there were many niches and rare genotypes sought out the niches in which they were best adapted, while common genotypes would also perforce have to occupy niches to which they were less well adapted. Wright and Dobzhansky [39] considered this possibility, and Wright found q for a simple case. Levene, Dobzhansky and Pavlovsky [15] and Lewontin [21] observed such cases and discussed its implications, and Haldane and Jayakar [8] considered it at more length. They showed that for genic selection, if the fitness of A_1 relative to A_2 is $1 - \phi(q)$, there is stable equilibrium with $q = Q$ if $\phi(Q) = 0$ and

$$0 < \phi'(Q) < \frac{2}{Q(1 - Q)}. \tag{4.4}$$

They point out that a model with overlapping generations gives rather different

results than the discrete generation model, since ϕ will then depend on the values of q at various times. This is unlike the cases we have discussed so far, where overlapping generations, or even a continuous model, makes little difference to equilibrium. Another new phenomenon here is that if $\phi'(Q)$ is very large, populations may overshoot the equilibrium value instead of approaching it asymptotically from one side, and may even go into undamped oscillations. Haldane and Jayakar find similar results for a recessive and for general selection in diploids with two alleles.

No mathematical analysis has been made for Ehrman's case of better mating ability of rarer genotypes, discussed in Dobzhansky's paper [3], but it should again lead to equilibrium.

5. The case of multiple loci and the optimum model

Recently much work has been done on the case of two or more loci, which may be linked, with recombination fraction $r \leqq 1/2$, with $r = 1/2$ if the loci are not linked. The interesting, and probably more important, case is where there are epistatic interactions on fitness; that is the fitness is not an additive function of the effects of the two loci separately. Here the poorly named phenomenon of "linkage disequilibrium" enters. We must work with the frequencies of gametes like A_iB_j and the frequency of such a gamete is no longer (freq A_i) (freq B_j).

Let us consider two loci A and B, each with two alleles, and let the alleles A_1, A_2, B_1, and B_2 have gene frequencies q, $1 - q$, p, $1 - p$, respectively. Under random mating, genotypes A_1A_1, A_1A_2, A_2A_2 will have frequencies q^2, $2q(1 - q)$, and $(1 - q)^2$ after one generation. However, if the initial gametic frequencies of A_1B_1, A_1B_2, A_2B_1, A_2B_2 are f_{11}^0, f_{12}^0, f_{21}^0, f_{22}^0, with $q = f_{11} + f_{12}$ and $p = f_{11} + f_{21}$, the equilibrium frequencies $f_{11} = qp$, and so forth, will be approached only asymptotically. At equilibrium

$$(5.1) \quad D = f_{11}f_{22} - f_{12}f_{21} = qp(1 - q)(1 - p) - q(1 - p)p(1 - q) = 0,$$

so that D is a measure of what is called linkage disequilibrium. If the crossover frequency r is the total frequency with which a heterozygote $(A_1B_1)/(A_2B_2)$ formed from the two gametes indicated forms the other two kinds of gametes (A_1B_2) and (A_2B_1) taken together, then $r = 1/2$ if the two loci are on different chromosomes and $r < 1/2$ if they are linked (on the same chromosome). Then, under random mating, the value of D in the tth generation is

$$(5.2) \quad D_t = (1 - r)^t D_0 \to 0 \quad \text{if} \quad r > 0.$$

This formula holds in the absence of selection. With selection, $D \to 0$, trivially if either locus reaches fixation, and also if each locus reaches equilibrium with additivity between the loci. If, on the other hand, there are epistatic interactions between the loci and either r is small enough or the interactions are large enough (even for $r = 1/2$), D will tend to a limit other than zero, and there is linkage disequilibrium. (See, for example, Lewontin and Kojima [26].)

Only the deterministic case has been considered in any detail, and except for particularly simple and symmetric sets of adaptive values, it has been necessary to resort to computers rather than analytic solutions. It would be easy to introduce a stochastic element by Monte Carlo runs, but the difficulty would be to do enough to obtain any real insight.

A particularly interesting case is the "optimum model" first formally developed by Wright [33]. Here, there is an underlying quantitative variable x that is influenced by several loci, each of which may involve several alleles. There is an optimum value G and the fitness of a genotype depends on its expected phenotype x by the relation $W = f(x)$, where $f(x)$ has a maximum f_0 at G and decreases monotonically to zero as x goes away from G. For x sufficiently close to G, such a function can be represented by $f_0 - c(x - G)^2$, and $W = f_0 - c(x - G)^2$ has been called the quadratic deviation optimum model. This model gives impossible negative W if x is too far from G. This can be avoided by using Malthusian parameters, and setting $m = F_0 - C(x - G)^2$. Then,

$$W = e^{F_0 - C(x-G)^2} = f_0 e^{-C(x-G)^2}, \tag{5.3}$$

which has the bell shape of a normal curve with its maximum equal to f_0 when $x = G$.

For any form of the optimum model, and for any reasonable model on the original x scale, there will be strong epistatic interactions. In particular, this will be so if the genetic effects on the x scale are linear and additive.

Now for the optimum model let us suppose that the differential effect of the A locus on the underlying character x is $-a$, Ha, a for A_1A_1, A_1A_2, and A_2A_2, respectively, and these numerical values are the same for each locus involved. There is no loss of generality in supposing that $H \geqq 0$. If $H = 0$, the heterozygote is exactly intermediate; if $H = 1$, there is complete dominance; and if $H > 1$, there is overdominance. It was shown by Wright [34] that there will be no equilibrium for the special cases $H = 0$ and $H = 1$, and hence, one would suppose, for intermediate values of H, but Kojima [13] showed that equilibrium is possible for other values of H, and that in the absence of linkage there are constants $C_1 > 0$ and $C_2 < 1$ such that there can be stable equilibrium if $C_1 < H < C_2$ and if the optimum value of x, G lies between limits depending on H. However, his results were partly incorrect because he assumed linkage equilibrium in his calculations. This has been a frequent stumbling block in work with more than one locus, since it greatly simplifies the algebra, but it must be justified in every case or errors result. Jain and Allard [10] allowed for linkage disequilibrium and found that in the absence of linkage slightly more stringent requirements were needed, with a smaller maximum value of G and with C_2 also smaller.

NOTE ADDED IN PROOF. Lewontin [40] shows that the reverse is true, and that allowing for linkage disequilibrium relaxes the requirements for gene equilibrium.

Lewontin [23] generalized Kojima's model to the special case of more than two loci given above. While he also neglected linkage disequilibrium, his conclusions are essentially correct. He pointed out that in order to keep L loci polymorphic, the optimum G has to be close to $G_1 + (2L - 1)a$ where the value of x is G_1 when all the loci are homozygous for the minus allele and $G_1 + 2La$ when they are all homozygous for the plus allele. Since selection keeps x near the optimum, it will be close to the maximum permitted by the genes present in the population and there will be little genetic variance in the character x.

In fact, Lewontin showed that the more loci are kept in stable equilibrium the *smaller* the total genetic variance will be. Furthermore, all the gene frequencies will be close to fixation of the plus allele if many alleles are being kept in equilibrium. Thus, there is opportunity for chance fixation at some loci; when this occurs frequencies of the plus alleles at the remaining loci can fall, and total variance go up, until ultimately a situation is reached with only one segregating locus, with both alleles quite common, and simple overdominance on the fitness scale.

If there is linkage between loci, polymorphism is easier to achieve and there can be more genetic variance. Polymorphism is also easier if there is overdominance on the x scale at some or all of the loci, or if there is a direct selective advantage to heterozygotes over and beyond that calculated from their x value. Perhaps the most obvious generalization of the Kojima-Lewontin model would be to have a constant a at all loci, but H at half and $-H$ at half, so that at half the loci the plus allele was partially dominant and at half the minus allele was partially dominant. This would allow an optimum near the midrange of x values. Even if it should turn out that gene frequencies were still near zero or one, such a situation would permit the population to respond to fairly large changes in the optimum value by changes in gene frequencies under natural selection.

It is probable that the specific mechanisms for maintaining genetic diversity within a population are inadequate, even if added together, to maintain the amount of diversity that is usually found in random mating populations. However, it is clear that the true situation involves all of them acting together, as well as others still unsuspected, and that their joint effects will probably be much greater than their simple sum. In particular the combination of numerous loci, interacting under the optimum model or by other mechanisms, with an environment varying both temporally and spatially in complex fashion, and with selective forces dependent on genotype frequencies, allows tremendous scope, both for genetic variation and for mathematical sophistication, even in the deterministic case.

REFERENCES

[1] S. KARLIN, J. MCGREGOR, and W. BODMER, "The rate of production of recombinants between linked genes in finite populations," *Proceedings of the Fifth Berkeley Symposium*

on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press, 1967, Vol. 4, pp. 403–414.

[2] E. R. DEMPSTER, "Maintenance of genetic heterogeneity," *Cold Spring Harbor Symp. Quant. Biol.*, Vol. 20 (1955), pp. 25–32.

[3] T. DOBZHANSKY, "Genetic diversity and diversity of environments," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 4, pp. 295–304.

[4] W. FELLER, "Diffusion processes in genetics," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1951, pp. 227–246.

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

[6] J. B. S. HALDANE, *The Causes of Evolution*, New York and London, Harper, 1932.

[7] J. B. S. HALDANE and S. D. JAYAKAR, "Polymorphism due to selection of varying direction," *J. Genet.*, Vol. 58 (1963), pp. 237–242.

[8] ———, "Polymorphism due to selection depending on the composition of a population," *J. Genet.*, Vol. 58 (1963), pp. 318–323.

[9] S. KARLIN and J. MCGREGOR, "The number of mutant forms maintained in a population," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 4, pp. 415–438.

[10] S. K. JAIN and R. W. ALLARD, "The nature and stability of equilibria under optimizing selection," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 54 (1965), pp. 1436–1443.

[11] M. KIMURA, "Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities," *Genetics*, Vol. 39 (1954), pp. 280–295.

[12] ———, "Diffusion models in population genetics," *J. Appl. Probability*, Vol. 1 (1964), pp. 172–232.

[13] K. KOJIMA, "Stable equilibria for the optimum model," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 45 (1959), pp. 989–993.

[14] H. LEVENE, "Genetic equilibrium when more than one ecological niche is available," *Amer. Natur.*, Vol. 87 (1953), pp. 331–333.

[15] H. LEVENE, O. PAVLOVSKY, and T. DOBZHANSKY, "Interaction of the adaptive values in polymorphic experimental populations of *Drosophila pseudoobscura*," *Evolution*, Vol. 8 (1954), pp. 335–349.

[16] R. LEVINS, "Theory of fitness in a heterogeneous environment. I. The fitness set and adaptive function," *Amer. Natur.*, Vol. 96 (1962), pp. 361–373.

[17] ———, "Theory of fitness in a heterogeneous environment. II. Developmental flexibility and niche selection," *Amer. Natur.*, Vol. 97 (1963), pp. 75–90.

[18] ———, "Theory of fitness in a heterogeneous environment. III. The response to selection," *J. Theoret. Biol.*, Vol. 7 (1964), pp. 224–240.

[19] ———, "The theory of fitness in a heterogeneous environment. IV. The adaptive significance of gene flow," *Evolution*, Vol. 18 (1965), pp. 635–638.

[20] ———, "Theory of fitness in a heterogeneous environment. V. Optimum genetic systems," *Genetics*, Vol. 52 (1965), pp. 891–904.

[21] R. C. LEWONTIN, "The effects of population density and composition on viability in *Drosophila melanogaster*," *Evolution*, Vol. 9 (1955), pp. 27–41.

[22] ———, "Evolution and the theory of games," *J. Theoret. Biol.*, Vol. 1 (1961), pp. 382–403.

[23] ———, "The interaction of selection and linkage. II. Optimum models," *Genetics*, Vol. 50 (1964), pp. 757–782.

[24] ———, "Selection in and of populations," *Ideas in Modern Biology* (edited by J. A. Moore), *Proceedings of the Sixteenth International Congress of Zoology*, Garden City, Natural History Press, Vol. 6, pp. 299–311.

[25] ———, "The genetics of complex systems," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 4, pp. 439–455.

[26] R. C. LEWONTIN and K. KOJIMA, "The evolutionary dynamics of complex polymorphisms," *Evolution*, Vol. 14 (1960), pp. 458–472.
[27] G. MALÉCOT, *Les Mathematiques de l'Heredite* (edited by Libraries de l'Academie de Medicine), Paris, Masson, 1948.
[28] S. P. H. MANDEL, "The stability of a multiple allelic system," *Heredity*, Vol. 13 (1959), pp. 289–302.
[29] P. A. P. MORAN, "The theory of some genetical effects of population subdivision," *Austral. J. Biol. Sci.*, Vol. 12 (1959), pp. 109–116.
[30] ———, *The Statistical Processes of Evolutionary Theory*, Oxford, Clarendon Press, 1962.
[31] E. POLLAK, "Stochastic theory of gene frequencies in subdivided populations," unpublished Ph.D. thesis, Columbia University, 1964. (University Microfilms, Ann Arbor, Michigan.)
[32] G. H. WEISS and M. KIMURA, "A mathematical analysis of the stepping stone model of genetic correlation," *J. Appl. Probability*, Vol. 2 (1965), pp. 129–149.
[33] S. WRIGHT, "Evolution in Mendelian populations," *Genetics*, Vol. 16 (1931), pp. 97–159.
[34] ———, "The analysis of variance and the correlations between relatives with respect to deviations from an optimum," *J. Genet.*, Vol. 30 (1935), pp. 243–256.
[35] ———, "Evolution in populations in approximate equilibrium," *J. Genet.*, Vol. 30 (1935), pp. 257–266.
[36] ———, "The differential equation of the distribution of gene frequencies," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 31 (1945), pp. 382–389.
[37] ———, "Isolation by distance under diverse systems of mating," *Genetics*, Vol. 31 (1946), pp. 39–59.
[38] ———, "Factor interaction and linkage in evolution," *Proc. Roy. Soc. Ser. B*, Vol. 162 (1965), pp. 80–104.
[39] S. WRIGHT and T. DOBZHANSKY, "Genetics of natural populations. XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila pseudoobscura*," *Genetics*, Vol. 31 (1946), pp. 125–156.
[40] R. C. LEWONTIN, "Stable equilibria under optimizing selection," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 56 (1966), pp. 1345–1348.