Proceedings of the National Academy of Sciences
Vol. 66, No. 1, pp. 87–93, May 1970

Selection and Polygenic Characters*

Montgomery Slatkin

PIERCE HALL, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS

Communicated by E. O. Wilson, February 20, 1970

Abstract. A mathematical model is developed which describes the effect of selection on polygenic or continuously varying phenotypic characters. For the simplest case the standard deviation of a phenotypic character in a population cannot be greater than $\sqrt{2}\beta$ where β is the expected deviation of the offspring from the average of the parents' types. This is found to carry over to density-dependent selection, and it indicates that there are limits to the amount of phenotypic diversity, and therefore niche breadth, which a population can maintain. These limits, however, are also subject to selection.

Introduction. A model will be presented here that describes the effects of normalizing selection on polygenic or continuously varying phenotypic characters. This type of selection is called either "optimizing"[1] or "stabilizing."[2] In the nuances of current usage, optimizing selection refers to any situation where an intermediate value of a phenotypic character is favored, while stabilizing selection is the artificial removal of extreme types in the course of selection experiments.[3]

Optimizing selection has been particularly difficult to model because of the gap in our knowledge between the well-founded theory of population genetics based on elementary Mendelian mechanics and empirical observations on the inheritance properties of polygenic characters. At this time we are unable to describe the effects of individual genes on a continuously varying character in order to develop an accurate model of the effect of selection. Ultimately, the effects of individual alleles will have to be known before a complete model can be made. Because of this present lack of knowledge, theoreticians are forced to make somewhat arbitrary assumptions about the effect of single alleles. They usually assume some additive value for each of the alleles or else prescribe nonadditive effects in a table that gives the correspondence between the possible genotypes and phenotypes. Models have been developed thereby which explore the consequences of selection, linkage, and inbreeding, either analytically with two loci or numerically with more loci.[4–6]

I would like to suggest a different and much more direct approach to this problem. Let us assume that there is a certain amount of heterozygosity present and then look at the consequences of different strengths of selection. We can proceed by assuming that the level of heterozygosity in a population manifests itself in the phenotypic diversity of the offspring of a single mating. Although the phenotypic diversity will of course depend on more than genetic factors, there must be some correlation between the phenotypes of the parents and offspring;

and it is this correlation by which the inheritance properties of the phenotypic character can be specified. In this way it is possible to avoid the necessity of a direct statement about the underlying genetic description of the population. This approach has the advantage of working directly at the phenotypic level where selection acts and the disadvantage of ignoring the genetic mechanisms which must be present. While the model will be necessarily incomplete, the principal result to be presented, that the phenotypic diversity is not strongly dependent on the strength of selection, should carry over to more accurate and detailed models.

The Model. Let us assume that the generations are nonoverlapping and that each generation consists of two parts, mating and the action of selection. The frequency distribution of a phenotypic character, z, is given by a function $p(z)$. $p(z)$ must be normalized, that is,

$$\int_{-\infty}^{\infty} p(z)dz = 1.$$

If z is outside physiological limits, then p is 0 so the limits of the integrals can always be infinite. Because of the artificial separation of selection and mating, there is a choice as to whether $p(z)$ describes the population before or after selection. Either choice will lead to essentially the same result, however, the mathematical formulation is slightly different. I will assume $p(z,t)$ is the distribution in generation t before selection.

The basic assumption for this model is that the distribution of the character in the offspring of a mating depends only on the value of character in the two parents. This can be described by a conditional probability, $L(z;z_1,z_2)$, which is the frequency distribution of z in the offspring of a mating of type z_1 with type z_2. L replaces the Mendelian formulation of mating. Since L is a probability, it also must be normalized.

$$\left(\int_{-\infty}^{\infty} L(z;z_1,z_2)dz = 1.\right)$$

In theory L could be determined from the careful analysis of breeding experiments, but this would be extremely difficult in practice.

The distribution in the next generation is found by summing over all possible combinations of matings, each weighted by its probability of occurrence. In a randomly mating population, the probability of type z_1 mating with type z_2 is $p(z_1)p(z_2)$. Thus, the distribution function in the next generation is given by

$$p(z,t+h) = \int_{-\infty}^{\infty} L(z;z_1,z_2)p_s(z_1,t)p_s(z_2,t)dz_1dz_2, \tag{1}$$

where h is the generation time. The subscript s is added on the right to emphasize that these are the distribution functions after selection, while $p(z,t)$ is the distribution before selection.

The action of selection is described by a function $S(z)$ which assigns a relative probability of survival to all individuals of type z. Since only the relative probabilities are important, S can be multiplied by a constant without changing the result. This is only true because population size is constant and the only question is which individuals make up the next generation. After selection, the distribution is changed to

$$p_s(z,t) = S(z)p(z,t)\Big/\int_{-\infty}^{\infty} p(z,t)S(z)dz, \tag{2}$$

where the denominator is present to assure that p_s is normalized. Let us assume that S does not change with time. From (1) and (2), the final equation for the model is

$$p(z,t+h) = \frac{\int_{-\infty}^{\infty}\int p(z_1,t)p(z_2,t)S(z_1)S(z_2)L(z;z_1,z_2)dz_1dz_2}{(\int_{-\infty}^{\infty} p(z,t)S(z)dz)^2}. \tag{3}$$

In principle, L and S can be experimentally determined and the time evolution of $p(z)$ can be measured and compared with the theoretical prediction to test the model. Let us be content to assume specific forms for L and S and explore the consequences of these assumptions.

The simplest form of L is found by assuming that z in the offspring of a mating of types z_1 and z_2 is distributed about the average of z_1 and z_2. Still a further simplifying assumption is that the variance of the offspring is independent of the parental types. It is unlikely that this could be accurate for all parental types. However the chance of a mating between two extreme types is small so an error in the description of their offspring will not affect the result by much. Therefore this assumption has to be satisfied in the range of z where most of the population is found. With these assumptions L can be written

$$L(z;z_1,z_2) = L\left(z - \frac{z_1+z_2}{2}\right), \tag{4}$$

where L is now a function of a single variable. We could reasonably expect L is symmetric about 0 and monotonically decreasing from its value at 0. L is characterized by a "width," β where

$$\beta^2 = \int_{-\infty}^{\infty} z^2 L(z)dz. \tag{5}$$

β is the expected deviation in the offspring from the average of the parental types. The basic assumption is that the heterozygosity in the population will manifest itself in β. There is no attempt to explain the maintenance of the heterozygosity. At the risk of increasing the semantic load in population genetics, let us call β the "reproductive variance." Even for a population which does not satisfy equation (4), it would be possible to define a measure of the reproductive variance which would reduce to (5) when (4) is satisfied. For exexample:

$$\beta^2 = \int_{-\infty}^{\infty}\left(z - \frac{z_1+z_2}{2}\right)^2 L(z;z_1,z_2)dz_1dz_2.$$

Now, of course, β depends on z_1 and z_2 but if this β is within certain bounds, then it still will be a useful quantity for describing the population. Let us show that β sets the scale of variability for a population.

At this time, we are interested in only the equilibrium situation. With (4), the equilibrium form of (3) is

$$p(z) = \frac{\int_{-\infty}^{\infty}\int L\left(z - \frac{z_1+z_2}{2}\right)p(z_1)p(z_2)S(z_1)S(z_2)dz_1dz_2}{(\int_{-\infty}^{\infty} p(z)S(z)dz)^2}. \tag{6}$$

Since β can change over a longer time scale, this is only a "quasi-equilibrium" as used in thermodynamics.

Optimizing selection: In the case of simple optimizing selection, $S(z)$ has a maximum at the optimum value, x, of the character. S then decreases from the maximum and has a characteristic width, σ, where

$$\sigma^2 = \int_{-\infty}^{\infty} (z - x)^2 S(z) dz. \tag{7}$$

σ is a measure of the strength of selection or the tolerance of the environment to nonoptimal phenotypes. Small σ represents strong selection and large σ weak selection.

When S is symmetric about x, we can expect the solution to equation (6), $p(z)$, to be a function with its mean at x and which is symmetric about $z = x$. If each side of (6) is multiplied by $(z - x)^2$ and integrated from $z = -\infty$ to $z = +\infty$, then

$$\alpha^2 = \beta^2 + \frac{\int_{-\infty}^{\infty} (z - x)^2 p(z) S(z) dz}{\int_{-\infty}^{\infty} p(z) S(z) dz}, \tag{8}$$

where

$$\alpha^2 = \int_{-\infty}^{\infty} (z - x)^2 p(z) dz. \tag{9}$$

α is the variance of the phenotypic character in the population. The second term on the right is positive so $\alpha^2 \geqslant \beta^2$. The largest value the second term can have is $\alpha^2/2$. This can be seen if we recall that

$$p_s(z) = \frac{p(z)S(z)}{\int_{-\infty}^{\infty} p(z)S(z)}$$

so that the integration is just giving the standard deviation of p_s, the distribution function after selection. Clearly, p_s could not have a larger standard derivation than p because optimizing selection could not increase the variance. Therefore

$$\beta^2 \leqslant \alpha^2 \leqslant 2\beta^2 \tag{10}$$

independent of the width of S. The lower limit is approached for small σ; the upper limit, for large σ. In the latter case, S does not change much in regions where p is nonzero. The scale of change of S is σ^2 and the scale of p is β^2. Therefore, if $\beta^2/\sigma^2 \ll 1$, in the integral in (8), S can be expanded in a power series about $z = x$. Thus, S can be replaced by

$$s(z) \simeq 1 + \frac{k}{\sigma^2}(z - x)^2 + 0\left(\frac{1}{\sigma^4}\right), \tag{11}$$

where k is a constant of $0(1)$. As long as selection is weak, any symmetric selection function is essentially the so-called "quadratic deviation" model. If selection is strong, the quadratic deviation model assigns negative fitnesses to a significant fraction of the population and a different functional form must be taken.

Density-dependent selection[11]**:** If, instead of simple optimizing selection, we consider the case of optimizing selection which is density dependent, the result does not change significantly. In this case, the selective advantage of type z depends on $p(z)$ as well as z. For example,

$$S(z) = S_0(z)(1 - \lambda p(z)), \tag{12}$$

where $S_0(z)$ is selective function for simple optimizing selection which was described above. The selective advantage of the optimal type is reduced as an increased fraction of the population has the optimal type. This could occur when z is a measure of some character which is adapted to a particular food size. If the food supply is limited, then the selection should be density dependent. λ describes the amount of competition coming from individuals with the same value of the phenotypic character.

We could expect that this type of selection should increase the variance of the character in the population. The question is by how much is α increased. While there is no simple analytic way to put bounds on α as there is for $\lambda = 0$, let us try to demonstrate that the upper bound on α is not changed by much for $\lambda > 0$, until λ becomes unreasonably large. This result is not surprising because, as $p(z)$ becomes flatter and broader because of the density effect, the importance of the density effect is decreased.

In the limit of very large σ, equation (8) becomes

$$\alpha^2 = \beta^2 + \frac{1}{2}\frac{\int_{-\infty}^{\infty} p(z)(1 - \lambda p(z))(z - x)^2 dz}{1 - \lambda \int_{-\infty}^{\infty} p^2(z) dz} \tag{13}$$

as long as $\alpha \ll \sigma$. The value of the integral term depends on the particular functional form of p, so the dependence of α on λ cannot be determined without more knowledge of the system. However, if we assume $p(z)$ is exactly Gaussian, then the integrals can be evaluated. This should provide an approximate result.

$$\alpha^2 = \beta^2 + \frac{1}{2}\frac{\alpha^2 - \lambda^2\sqrt{\pi}/2}{1 - \lambda\sqrt{\pi}/2\alpha}. \tag{14}$$

The positive solution to this cubic equation is given by standard formulas as

$$\alpha = \sqrt{\frac{a}{3}} \cos \phi, \tag{15}$$

where $a = 2\beta^2 + (3\pi/16)\lambda^2$ and $0 \leqslant \phi \leqslant \pi/6$. Thus α is bounded by $\sqrt{2\beta^2}[1 + (3\pi/3^2)(\lambda^2/\beta^2)]^{1/2}$. Of course this is only an approximation because p is not necessarily Gaussian. This model was tested numerically by assuming a Gaussian form for L and S and finding the solution to equation (6) with S given by (12). This was done by iterating equation (3) for different initial distributions until equilibrium was reached. This also confirmed the fact that the equilibrium solution is stable and independent of initial conditions. The results are shown in Figure 1 for $\lambda = 0, 2, 4$. In Table 1, the standard deviations found from the computer study are compared with those estimated from equation (14). The second values are consistently larger indicating that the actual distribution is flatter than the Gaussian with the same standard deviation. That is, the higher order moments are larger than those for the Gaussian. Thus, equation (15) provides the upper bound to α as σ increases, although this would have no meaning if λ is too large. However, for moderate values of λ and σ, it is clear that α will still be of the same order of magnitude as β and the diversity is still limited by the reproductive variance.

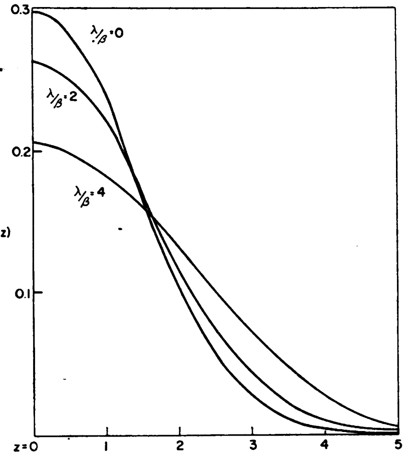

FIG. 1.

TABLE 1.

λ/β	α/β Result from computer	α/β Approx. result using eq. (14)
1	1.435	1.76
2	1.543	2.16
3	1.701	2.74
4	1.901	3.38
5	2.107	...

Discussion. The consequence of this simple model is that the variance of a phenotypic character in a population is primarily determined by the reproductive variance and only secondarily by the strength of selection. The exact value of α, the variance of a character, depends on the particular form of the selection function, $S(z)$, but the range of possible values is limited. The significance of this result is that it is not necessary to assume the existence of strong selective forces in order to explain a lack of phenotypic diversity. Even in the presence of weak selection, there is no tendency for the distribution function, $p(z)$, to broaden beyond $\sqrt{2}\ \beta$. This is because of the random mating which causes extreme types to mate most frequently with types near the mean, causing their offspring to be less extreme.

The diversity of certain phenotypic characters has often been taken by ecologists as a measure of niche breadth, on the assumption that the values of the phenotypic character places a limit on the part of the environment which can be utilized.[6, 7] For such a character it is commonly assumed that the niche breadth is limited by competition from other population. But the present analysis shows that the genetic system places a stringent upper limit on niche breadth. This could explain why, in the absence of competition, many island populations still show little or no increase in phenotypic variance. Moreover, the increase in variance observed could be accounted for by an increase in the density effect. If the food supply is lower on an island, then there is increased competition from individuals of the same type, so λ is larger.

The function $L(z;z_1,z_2)$, which describes the inheritance properties of a character, must also be thought of as being affected by selection, only over a longer time scale. For example, if $S(z)$ is a narrow ($\sigma \ll \beta$), there would be a selective advantage to reducing the reproductive variance in order to reduce the segregational load. This would not be group selection because an individual which tends to have its offspring nearer the optimal value would individually have a higher fitness. Hence, there would be a long term selective advantage to decreasing β. This would have to be balanced by a selective force to increase β, that is, a selective force for increased heterozygosity. This would also act over a long time and would be a consequence of the need to adapt to changes in the environment. Mather[8] has called this balance of selective forces the "relational balance" which determines the genetic background of the character. Unless the population is subject to violent changes in environmental conditions it is reasonable to suppose that selection would act on L until $\beta \simeq \alpha$. The advantage of reducing β beyond this would not be significant.

Biologists have recognized that natural selection can act over different time scales at the same time. Waddington[9] emphasized this by the use of the terms "normalizing" and "canalizing" selection. In terms of this model, $S(z)$ describes the normalizing selection, while the selection acting to alter L is the canalizing selection. Mayr[10] discusses this same idea in terms of the effect of cohesive factors (in the genetic system) on selection. He points out that cohesive factors are also subject to selective forces. Most mathematical modelers have concentrated their efforts on describing the effects of normalizing selection, without accounting for the fact that selection can act to modify the effect of normalizing forces. It is clear from the results of this model that both normalizing and canalizing selection must be accounted for before a full description of the evolution of a population can be accomplished.

* I would like to thank Professors W. H. Bossert and E. O. Wilson of Harvard University for their help and encouragement during the preparation of this paper. This work was supported in part by the National Science Foundation under grant NSF:GP-8666 and by the Division of Engineering and Applied Physics, Harvard University.

[1] Lewontin, R. C., *Genetics*, **50**, 757–782 (1964).

[2] Mather, K., *Symp. Soc. Exp. Biology*, **7**, 66–95 (1953).

[3] Thoday, J. M., *Heredity*, **13**, 187–203 (1959).

[4] Kimura, M., *Evolution*, **10**, 278–287 (1956).

[5] Lewontin, R. C., *Genetics*, **49**, 49–67 (1964).

[6] Levins, Richard, *Evolution in a Changing Environment* (Princeton, N.J.: Princeton University Press, 1968).

[7] Van Valen, L., *Amer. Naturalist*, **99**, 377–390 (1965).

[8] Mather, K., *J. Genetics*, **41**, 159–193 (1941).

[9] Waddington, C. H., *The Strategy of the Gene* (London: George Allen and Unwin, 1957).

[10] Mayr, E., *Animal Species and Evolution* (Cambridge: Harvard U. Press, 1963).

[11] *Note added in proof:* It has been suggested that "frequency-dependent selection" is a better term than "density-dependent selection" $S(z)$ depends on the frequency of type z, not on the over-all population density.