THE EVOLUTION OF NON-ADDITIVE GENETIC VARIANCE UNDER ARTIFICIAL SELECTION

I. MODIFICATION OF DOMINANCE AT A SINGLE AUTOSOMAL LOCUS

By B. D. H. Latter*

[Manuscript received November 25, 1963]

Summary

The papers to be published in this series are concerned with the behaviour under directional selection of genetic systems showing non-allelic interaction. The aim is to describe in quantitative terms the conditions under which overdominance may evolve under the usual conditions of laboratory experimentation, and to examine the stability of the resultant populations. In this first paper, a general method of approach to problems of the modification of dominance is outlined, and the necessary theory developed.

I. INTRODUCTION

The long-term effects of artificial selection on the behaviour and genetic structure of a population can at present be described in only the most general terms. A. Robertson (1960) has examined the effects of restricted population size on the ultimate limits to selection, for a character controlled by independent genes of small effect, and has shown the expected total advance in a given population to be a function only of the effective population size and the intensity of selection. Kojima (1959) has set out the conditions under which stable equilibria are possible in a population undergoing mass selection, ignoring the complications due to linkage disequilibrium; and Lewontin and Kojima (1960) have studied the joint effects of linkage and epistasis in the approach to equilibrium of a two-locus system under selection.

Experimental studies have highlighted the complexity of the genetic situation prevailing in populations which have reached a plateau under selection. The analyses reviewed by F. W. Robertson (1955) of the effects of artificial selection for increased body size in *Drosophila*, for example, have identified three distinct phenomena which may be responsible for the lack of response in populations at a plateau: (i) the loss of genes due to restricted population size; (ii) the antagonistic effects of natural selection; and (iii) effective overdominance, leading to the retention of genetic variance under long-continued selection. Evidence to the same effect has been provided by Mather and Harrison (1949), Dickerson (1955), Falconer (1955), Clayton and A. Robertson (1957), and by Brown and Bell (1961).

The observed residuum of non-additive genetic variance in a population under directional selection can be accounted for in a number of ways. It is, of course,

* Division of Plant Industry, CSIRO, Canberra.

entirely possible that overdominance for the character is displayed at a minority of loci in the unselacted population, and that the effects of segregation at these loci are simply accentuated during selection by the attainment of intermediate gene frequencies at equilibrium, or by changes in the genetic background. Alternatively, chance loss of gametes may lead to the effective overdominance of a chromosome segment, if it carries closely linked loci showing partial dominance.

Wright (1935) has suggested that overdominance may in fact evolve, under selection, from an initially additive genetic system, if maximum performance depends on the realization of optimal mean values for underlying processes. It is well known that continued selection for such optima leads to the elimination of genetic variability, but it is unlikely that this process could proceed to appreciable extent in the time-scale of a laboratory experiment (A. Robertson 1956).

In the present series of papers, it is proposed to study the behaviour under directional selection of a number of genetic models, beginning with the simplest and working through to the more complex. The aim is to describe in quantitative terms the conditions under which overdominance may evolve under the usual conditions of laboratory experimentation, and to examine the stability of the resultant genetic systems. In this first communication, a general method of approach to problems of the modification of dominance is outlined, and the necessary theory elaborated.

II. The Theory of a Single Locus Segregating against an Arbitrary Genetic Background

The approximate effects of selection on the genotypes at a single segregating locus, irrespective of the complexity of their interaction with the remainder of the genotype, can be examined by means of a device analogous to that proposed by Falconer (1952) for the study of genotype × environment interaction. The population can be considered to be composed of a number of subpopulations, one corresponding to each of the genotypes at the locus under scrutiny: one can then specify the additive genetic variance present in each subpopulation, and the additive genetic correlation between the performance of background genotypes in each pair of subpopulations. In this section, the consequences of artificial selection operating on an arbitrary genetic system which may show dominance and epistatic effects will be set out in terms of these parameters. The behaviour of the population will be described from the viewpoint of the modification of the average values of the genotypes at a single autosomal locus with two alleles: the complications due to linkage relationships among the loci in the system will not be considered.

Suppose initially that the genetic system involves three loci A, B, and C, each with two alleles at frequencies p_1, p_2; q_1, q_2; r_1, r_2. Denote the average performance of the genotype A_1A_2B_mC_mC_m by y_{121212}. The total additive genetic variance contributed by the three loci can then be shown to be

$$\sigma_A^2 = \frac{1}{2} p_1 p_2 \left( \frac{\partial \bar{y}}{\partial p_1} \right)^2 + \frac{1}{2} q_1 q_2 \left( \frac{\partial \bar{y}}{\partial q_1} \right)^2 + \frac{1}{2} r_1 r_2 \left( \frac{\partial \bar{y}}{\partial r_1} \right)^2,$$

(1)

where $\bar{y}$ is the mean of the population.
Let the average value of a genotype at the first locus, i.e. \( y_{ij} \), be denoted simply by \( y_{ij} \): then the partial derivative of \( \bar{y} \) with respect to \( q_1 \) can be written

\[
\frac{\partial \bar{y}}{\partial q_1} = \frac{\partial}{\partial q_1}(q_1^2 y_{ij} + 2p_1 p_2 y_{12} + p_2^2 y_{22})
\]

and similarly

\[
\frac{\partial \bar{y}}{\partial r_1} = \frac{\partial}{\partial r_1}(p_1^2 \frac{\partial y_{11}}{\partial q_1} + 2p_1 p_2 \frac{\partial y_{12}}{\partial q_1} + p_2^2 \frac{\partial y_{22}}{\partial q_1}),
\]

so that if \( A \) is taken to be the primary locus under observation, and loci \( B \) and \( C \) provide the genetic background variation, the total additive genetic variance can be rewritten as

\[
\sigma_A^2 = \frac{1}{2} p_1 p_2 \left( \frac{\partial \bar{y}}{\partial q_1} \right)^2 + p_1^2 \left[ \frac{\partial y_{11}}{\partial q_1} \right]^2 + \frac{1}{2} r_1 \left( \frac{\partial y_{11}}{\partial r_1} \right)^2
\]

\[
+ 4p_1^2 p_2 \left[ \frac{\partial y_{11}}{\partial q_1} \right]^2 \left( \frac{\partial y_{12}}{\partial q_1} \right)^2 + \frac{1}{2} r_1 \left( \frac{\partial y_{11}}{\partial r_1} \right) \left( \frac{\partial y_{12}}{\partial r_1} \right)
\]

\[
+ 4p_1^2 p_2 \left[ \frac{\partial y_{11}}{\partial q_1} \right] \left( \frac{\partial y_{12}}{\partial q_1} \right) \left( \frac{\partial y_{22}}{\partial q_1} \right) + \frac{1}{2} r_1 \left( \frac{\partial y_{11}}{\partial r_1} \right) \left( \frac{\partial y_{22}}{\partial r_1} \right)
\]

\[
+ 2p_1 p_2 \left[ \frac{\partial y_{11}}{\partial q_1} \right] \left( \frac{\partial y_{22}}{\partial q_1} \right) + \frac{1}{2} r_1 \left( \frac{\partial y_{11}}{\partial r_1} \right) \left( \frac{\partial y_{22}}{\partial r_1} \right)
\]

\[
= \sigma_{A(1)}^2 + p_1^2 \sigma_{A(11)}^2 + 4p_1^2 p_2^2 \sigma_{A(12)}^2 + p_2^2 \sigma_{A(22)}^2 + 4p_1^2 p_2^2 \text{cov}_{A(11, 12)} + 4p_1 p_2^2 \text{cov}_{A(12, 22)}
\]

(2)

where

\( \sigma_{A(1)}^2 \) is the additive genetic variance contributed by the primary locus;

\( \sigma_{A(11)}^2 \) is the additive genetic variance contributed by the background loci in the presence of \( A_i A_j \);

and

\( \text{cov}_{A(11, 12)} \) is the additive genetic covariance between the performance of background genotypes in the presence of \( A_i A_j \) and \( A_i A_k \).

Ignoring complications due to the generation of linkage disequilibrium between loci in the system, the response in the mean of the population following one generation of artificial selection is

\[
\Delta \bar{y} = (\partial \bar{y} / \partial q_1) \Delta q_1 + (\partial \bar{y} / \partial r_1) \Delta r_1.
\]

Suppose the total variance in the quantitative character under selection to be \( \sigma^2_A \); then according to the usual theory based on genes of individually small effect,

\[
\Delta p_1 = \frac{1}{2} p_1 p_2 \left( \frac{\partial \bar{y}}{\partial q_1} \right);
\]

\[
\Delta q_1 = \frac{1}{2} q_1 q_2 \left( \frac{\partial \bar{y}}{\partial q_1} \right);
\]
\[ \Delta r_1 = i \frac{\mathcal{R}_1}{2\sigma^2} \] 

so that

\[ \Delta \mathcal{R} = \left( \frac{i}{\sigma_p} \right) \sigma_j^2. \]  

By means of equation (2) this total advance may be analysed into the following components:

1. The response due to the change in gene frequency at the primary locus, i.e. \( \left( \frac{i}{\sigma_p} \right) \sigma_j^2 \); and

2. The response due to changes in the absolute magnitudes of the \( y_{11} \), i.e.

\[ p_1^1 \Delta y_{11} + 2p_1p_2 \Delta y_{12} + p_2^2 \Delta y_{22}, \]

where

\[ \Delta y_{11} = \left( \frac{i}{\sigma_p} \right) \left[ p_1^1 \sigma_j^2(11) + 2p_1p_2 \text{COV}_{A(11,12)} + p_2^2 \text{COV}_{A(11,22)} \right], \]

\[ \Delta y_{12} = \left( \frac{i}{\sigma_p} \right) \left[ p_1^1 \sigma_j^2(11,12) + 2p_1p_2 \sigma_j^2(12) + p_2^2 \text{COV}_{A(11,12)} \right], \]

and

\[ \Delta y_{22} = \left( \frac{i}{\sigma_p} \right) \left[ p_2^2 \sigma_j^2(12,22) + 2p_1p_2 \text{COV}_{A(11,22)} + p_2^2 \sigma_j^2(22) \right]. \]

In addition to its effects on the mean of the population, selection will in general lead to changes in the relative magnitudes of the \( y_{ij} \); these changes depend on five parameters which can most usefully be specified as follows:

\[
\begin{align*}
\alpha &= \frac{1}{3} \sigma_j^2(11) + \frac{1}{3} \sigma_j^2(22) - \text{COV}_A(11,22); \\
\beta &= \sigma_j^2(12) - \frac{1}{2} \sigma_j^2(11) - \frac{1}{2} \sigma_j^2(22); \\
\gamma &= \sigma_j^2(22) - \sigma_j^2(11); \\
\delta &= \sigma_j^2(12) - \frac{1}{2} \text{COV}_A(11,12) - \frac{1}{2} \text{COV}_A(12,22); \\
e &= \text{COV}_A(12,22) - \text{COV}_A(11,12).
\end{align*}
\]

The change in the relative positions of the two homozygotes at the primary locus, i.e. the "proportionate effect" of the gene (Falconer 1960), is in general given by

\[ \Delta (y_{22} - y_{11}) = \left( \frac{i}{\sigma_p} \right) \left[ p_1^1 \left( \frac{1}{2} \gamma - \alpha \right) + 2p_1p_2(\epsilon) + p_2^2 \left( \frac{1}{2} \gamma + \alpha \right) \right], \]  

(4a)

and the change in the position of the heterozygote relative to the mean of the two homozygotes is

\[ \Delta (y_{12} - y_{11} - y_{22}) = \left( \frac{i}{\sigma_p} \right) \left[ p_1^1 \left( \frac{1}{2} \alpha + \beta + \frac{1}{2} \gamma - \delta - \frac{1}{2} \epsilon \right) + 2p_1p_2(\delta) + p_2^2 \left( \frac{1}{2} \alpha + \beta - \frac{1}{2} \gamma - \delta + \frac{1}{2} \epsilon \right) \right]. \]  

(4b)

In the case of a gene segregating independently of the genetic background, each of the five parameters is zero, and both the proportionate effect and degree of dominance of the locus are unaffected by selection.

Although the above theory has been developed for only two background loci, it is quite clear that it can be extended to deal with any number of such loci, provided they have individually small effects. It is also obvious that any one of the segregating loci can be taken to be the primary locus, and the remainder as providing the genetic background. However, the theory can be expected to give useful predictions only over a limited number of generations; for it obviously cannot take account of changes in the parameters which specify the nature of the genetic interaction between primary locus and background.
III. The Modification of Dominance

Despite the short-term nature of the theory which has been outlined, it is profitable to consider in some detail the conditions which favour the progressive increase of dominance at a given locus. Let \( d = (y_{12} - \frac{1}{2}y_{11} - \frac{1}{2}y_{22}) \) and \( a = (\frac{1}{2}y_{22} - \frac{1}{2}y_{11}) \), so that \( d/a \) measures the average degree of dominance displayed at the locus. Under artificial selection, the change in \( d/a \) in one generation is

\[
\Delta(d/a) = (d+\Delta d)/(a+\Delta a) - d/a = \frac{\Delta d - (d/a)\Delta a}{(a+\Delta a)}, \tag{5}
\]

provided \( a \neq 0 \). The following discussion will be restricted to a consideration of loci for which \( a > 0 \), and \( \Delta a > -a \), so that the denominator in equation (5) is positive.

It is also useful to define measures of the additive genetic variance displayed by each of these properties of the primary locus, and of the covariance between them, as the background genotype is varied in the population. In terms of the five parameters previously defined, it turns out that

\[
\begin{align*}
\var_A(a) &= \frac{1}{2}a, \\
\var_A(d) &= 2\delta - \beta - \frac{1}{2}a, \\
\text{and} \quad \cov_A(a, d) &= \frac{1}{2}(\epsilon - \frac{1}{2}y).
\end{align*}
\]

(a) Initial Change at an Additive Locus

If \( d/a = 0 \), the condition that \( \Delta(d/a) \) should be positive under artificial selection for increased expression of the character is simply given by

\[
(p_1^2 + p_2^2)(\beta + \frac{1}{2}a) - (p_1 - p_2)^2 \delta + \frac{1}{2}(p_1 - p_2)(\epsilon - \frac{1}{2}y) > 0
\]

from (4b) and (5). Substituting \( p_2 = \frac{1}{2} + p \), so that \( -\frac{1}{2} < p < +\frac{1}{2} \), the inequality becomes

\[
-2p^2(2\delta - \beta - \frac{1}{2}a) + p(\epsilon - \frac{1}{2}y) + \frac{1}{2}(\beta + \frac{1}{2}a) > 0.
\]

If the values of the parameters are such that this inequality is satisfied for \( p = \pm \frac{1}{2} \), it will also be satisfied for any other legitimate value of \( p \). The condition that dominance at an additive locus should begin to evolve immediately in the direction of artificial selection, irrespective of the initial gene frequency, is then given by

\[
\beta + \frac{1}{2}a - \delta > \frac{1}{2} | \epsilon - \frac{1}{2}y | . \tag{6}
\]

Since \( 2\delta - \beta - \frac{1}{2}a > 0 \), it follows that if (6) is satisfied, \( \delta \) is positive and hence \( \beta + \frac{1}{2}a \) is also positive.

The magnitude of \( \Delta d \) will, in general, be related to that of the integral

\[
\int_{-\frac{1}{2}}^{+\frac{1}{2}} [-2p^2(2\delta - \beta - \frac{1}{2}a) + p(\epsilon - \frac{1}{2}y) + \frac{1}{2}(\beta + \frac{1}{2}a)] dp = \frac{1}{2}(2\beta + a - \delta),
\]

and will therefore be expected to be a maximum when

(i) the additive genetic variance in the presence of the heterozygote, \( \sigma^2_{A[H]} \), exceeds the mean of the corresponding variances in the presence of the homozygotes; and/or
(ii) the additive genetic correlation between performance in the presence of the two homozygotes, \( r_{A(11,22)} \), is less than the mean of the corresponding correlations involving the heterozygote.

The change in the degree of dominance, \( d/a \), will also depend on \( \Delta a \): if \( \Delta a \) is small in absolute magnitude by comparison with \( a \), the change in \( d/a \) will be greatest when \( \Delta a \) is a minimum, i.e. when

\[
\int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ -p^2(e-\frac{1}{3}y) + p + \frac{1}{2}(e+\frac{1}{3}y) \right] dp = \frac{1}{3}(e+\gamma)
\]

is a minimum.

(b) Conditions Favouring the Evolution of Overdominance

Under continued artificial selection, the evolution of overdominance will be favoured if the values of the parameters are such that \( \Delta d-(d/a)\Delta a \) is positive, irrespective of the value of \( p \) in the range \(-\frac{1}{2} \leq p \leq \frac{1}{2} \), or of the value of \( d/a \) in the range \( 0 < d/a < 1 \). Integration with respect to \( p \) and \( d/a \) gives

\[
\int_{0}^{1} \int_{-\frac{1}{2}}^{\frac{1}{2}} [\Delta d-(d/a)\Delta a] dp \cdot d(d/a) = \frac{1}{3}(i\sigma_p)[(2\beta+a-8)-\frac{1}{3}(e+\gamma)]. \quad (7a)
\]

The general conditions that the change in \( d/a \) should be a maximum under continued selection are therefore that

\[
[s_{A(11,22)}^2 - \text{cov}_A(11,22)] - \frac{1}{4}[s_{A(11)}^2 - \text{cov}_A(11,12)] - \frac{1}{4}[s_{A(22)}^2 - \text{cov}_A(11,12)] \quad (7b)
\]

be maximized.

As before we can distinguish between genetic interaction determined by the scale of the additive genetic effects displayed in the presence of each genotype at the primary locus, and that due to imperfect additive genetic correlation between performance in the presence of alternative genotypes at the locus. Let us consider firstly the situation of homogeneous variances specified by \( \beta = \gamma = 0 \), in which the genotypes at the primary locus have no influence on the scale of the additive genetic effects due to the residual genotype. Suppose further that \( \epsilon = 0 \), so that the pattern of interaction can be specified by two correlation coefficients, viz.

\[
r^* = r_{A(11,22)}
\]

and

\[
r = r_{A(11,12)} = r_{A(12,22)}.
\]

Inequality (6) then requires that \( a-2\delta \) be positive for an initial increase in the degree of dominance at an additive locus, i.e. that

\[
(1-r^*) > 2(1-r). \quad (8)
\]

We must also take account of the usual restriction that the determinant of the correlation matrix be non-negative, i.e. that

\[
1+2r^*r^2-r^*^2-2r^2 \geq 0.
\]

It can readily be shown that this expression is zero when either \( r^* = 1 \) or \( r^* = 2r^2-1 \), and that for a given value of \( r \) the inequality is satisfied by any \( r^* \) lying within this interval. Hence we may write

\[
(1-r^*) \leq 2(1-r^2). \quad (9)
\]
Figure 1 shows the range of values of $r^*$ satisfying equations (8) and (9), as a function of the value of $r$.

In addition to the requirement that equation (6) be satisfied for an initial increase in dominance at an additive locus, equation (7) indicates that $\alpha-\delta$ should be large if a further rapid increase in the degree of dominance is to be effected by directional selection. Note, however, that $\delta$ is non-negative in the present context, so that the value of $\alpha$ must therefore be large. Reference to equation (4a) shows that this can be expected to lead to marked changes in the proportionate effect of the gene concerned, particularly at extreme gene frequencies. When the desirable allele ($A_2$) is at low frequency, a reduction in the proportionate effect is expected under selection, whereas when $p_2$ is high the difference in average value between the two homozygotes will increase.

At the other extreme, consider the situation in which $r_{A(11,12)} = r_{A(12,22)} = r_{A(11,22)} = 1$, the genetic interaction between primary locus and background being due solely to an increase in the scale of additive genetic effects in the presence of the heterozygote at the primary locus (i.e., $\beta > 0$). If the two homozygotes have the same variance, it follows that $\gamma = \epsilon = \alpha = 0$, and that

$$\beta = \frac{\sigma_{A(12)} + \sigma_{A(22)}}{\sigma_{A(12)} - \sigma_{A(22)}}$$

and

$$\delta = \sigma_{A(12)}/\sigma_{A(22)}.$$
so that $\beta > 0$ ensures that equation (6) is satisfied. Equation (4a) then shows the proportionate effect of the gene to be unaltered by selection under these conditions. If the additive genetic variance in the presence of the favoured homozygote, $\sigma_a^2(21)$, exceeds that in the presence of the alternative homozygote, then

$$\gamma = \sigma^2_a(21) - \sigma^2_a(11),$$

$$\epsilon = \sigma^2_a(12) - \sigma^2_a(11),$$

and

$$\frac{1}{2}y - a = \sigma^2_a(11) - \sigma^2_a(12),$$

are all positive, and the change in the proportionate effect of the locus is shown by equation (4a) to be positive, irrespective of the prevailing gene frequency. If $\sigma^2_a(22) < \sigma^2_a(11)$, then $\gamma$, $\epsilon$, and $\frac{1}{2}y - a$ are negative, and

$$\frac{1}{2}y + a = \sigma^2_a(21) - \sigma^2_a(11),$$

is also negative. The change in the magnitude of the proportionate effect of the gene will consequently be negative for any value of $p_2$.

IV. DISCUSSION

The theory developed in this paper can be considered to give no more than a first approximation to the precise changes in dominance expected under long-continued artificial selection. It has already been emphasized that the theory holds strictly only for genes of small effect, and that it ignores complications due to the generation of linkage disequilibrium. In addition, the five basic parameters specifying the nature of the genetic interaction between the primary locus and the background genotype must be considered to remain unchanged in any application of the theory. One might therefore expect long-term changes to be exaggerated by comparison with those in a finite genetic system.

Nevertheless, the approach can profitably be utilized in a numerical study of the conditions under which overdominance will evolve under continued directional selection. In such a study one can drop the restriction that the primary locus should have a small effect on both the mean and the residual variance of the character, provided a suitable definition of selective values under artificial selection is adopted. It is proposed to give the results of such a survey in the next paper in this series. Tentative conclusions reached in this way can then provide a guide to the construction of finite genetic models, whose behaviour under selection can with profit be studied by exact numerical methods.

V. References


EVOLUTION OF NON-ADDITIVE GENETIC VARIANCE. I


FALCONER, D. S. (1960).—“Introduction to Quantitative Genetics.” (Oliver and Boyd: Edinburgh.)


