

ON MODELS OF QUANTITATIVE GENETIC  
VARIABILITY: A STABILIZING  
SELECTION-BALANCE MODEL

*Lev A. Zhivotovsky and Marcus W. Feldman*

*Paper number 0038*

Institute of General Genetics  
USSR Academy of Sciences  
Gubkin St. 3, 117 809 GSP-1  
Moscow B-333, USSR

Department of Biological Sciences  
Stanford University  
Stanford, California 94305  
Stanford University

1991



## **Abstract**

A model of stabilizing selection on a multilocus character is proposed that allows the maintenance of stable allelic polymorphism and linkage disequilibrium. The analysis is carried out for weak selection in a quadratic-deviation model for the stabilizing selection. The stationary state is characterized by unequal allele frequencies, unequal proportions of complementary gametes, and a reduction of the genetic (and phenotypic) variance by the linkage disequilibrium. The model is compared with Mather's polygene balance theory, with models that include mutation-selection balance, and others that have been proposed to study the role of linkage disequilibrium in quantitative inheritance.



## Introduction

The population genetic description of quantitative variation in natural populations has long been of major interest to natural historians (see e.g. Futuyma, 1986, ch 7). Besides the development of an appropriate statistical framework to discuss such variation, the questions of its maintenance and change under natural selection have received much attention in the theoretical literature (e.g. Lande, 1975; Turelli and Barton, 1990).

Fisher (1918) proposed a variance decomposition for a quantitative trait determined by a set of polymorphic loci in the absence of selection. There have been a number of extensions to the original theory that involve assumptions about the absence of the linkage disequilibrium and interactions between components of the variation. One of them has the form (Cockerham, 1954; Kempthorne, 1957)

$$V = V_A + V_I + V_E \quad (1)$$

where  $V$  is the overall phenotypic variance,  $V_A$  is the additive genetic variance, and  $V_I$  is a component due to interactions between alleles and among genes.  $V_E$  is a measure of (non-transmitted) environmental variance. The genotypic components  $V_A$  and  $V_I$  are determined by a set of constant parameters as well as the frequencies of alleles (see Falconer, 1989, for the details). If, as is often assumed to be the case, there is stabilizing selection on the trait under study, then it becomes important to understand how allelic and genotypic frequencies change in response to the selection on the phenotype, and how these changes affect the decomposition (1).

In one of the earliest attempts to address genotypic response to phenotypic selection, Mather (1942) developed a model of *polygene balance*. According to this model “polygenes” are distributed along chromosomes in such a way that the signs of their contributions to the phenotype alternate. Thus, polygenes with larger (+) and smaller (–) effects on the trait follow each other in the sequence + – +– etc. Mather’s model also supposes the presence of complementary pairs of chromosomes whose products produce the genotypic balance: + – + – ⋯ / – + – + ⋯. This polygenic balance may be invoked to explain the greater fitness under stabilizing selection of individuals with such complementary pairs of chromosomes relative to those in which configurations like – – – – ⋯ or + + + + ⋯

predominate. Obviously, Mather's model allows production of less balanced polygene combinations by recombination, but in his view the balanced complementary pairs are the primary contributors to quantitative variability under stabilizing selection.

In the terminology of population genetics, Mather's idea involves the presence of linkage disequilibrium, since it involves complementary pairs of gametes. Mather's treatment did not lead to an explicit computation of linkage disequilibrium produced by selection. It is well known, however, that selection can lead to stable linkage disequilibrium (Lewontin and Kojima, 1961; Bodmer and Felsenstein, 1967; Karlin and Feldman, 1970; Franklin and Lewontin, 1970; Feldman et al., 1974). Bulmer (1974, 1980) studied model (1) under stabilizing selection and concluded that it results in linkage disequilibrium. The arguments were not, however, structured as formal population genetic models. In particular he used a completely additive variant of model (1) that actually has no stable polymorphic state. Lande (1975) and Turelli and Barton (1990) recognized this instability and used a mutation-selection balance to maintain polymorphic variation. Within this framework linkage disequilibrium is not an important contributor to the ultimate distribution of the quantitative character.

Stable polymorphic states may exist under epistatic selection (Nagylaki, 1989; Gillespie and Turelli, 1989) and may exhibit linkage disequilibrium (see references above). It may therefore be the case that alternative models of *stabilizing selection* may (a) be appropriate, and (b) exhibit stable polymorphic equilibria with linkage disequilibrium. Some models of this kind have been discussed by Zhivotovsky and Gavrilets (1992) and in this paper we return to these models. The nature of the stable polymorphic states that we are able to study, under stabilizing selection, suggests that Mather's polygene balance is not a tenable hypothesis, that Bulmer's model is inaccurate in computing the change of quantitative variability under stabilizing selection and that linkage disequilibrium may be important to the variation of quantitative characters at the stationary state under stabilizing selection.

## The Model

Consider a quantitative character affected by  $n$  autosomal loci each with two alleles

which we denote by  $A_i$  and  $a_i$ . The contribution of the  $i$ -th locus to the character is measured by a weight  $\alpha_i$  ( $\alpha_i > 0$ ). The relative contributions of the alleles  $A_i$  and  $a_i$  at the  $i$ -th locus to the character are 1 and 0 respectively. Let  $G = (\ell_1, \ell_2, \dots, \ell_n)$  denote a gamete where  $\ell_i$  are indicator variables such that  $\ell_i = 1$  if the gamete contains  $A_i$  and  $\ell_i = 0$  if the gamete contains  $a_i$  at the  $i$ -th position. If  $G' = (\ell'_1, \ell'_2, \dots, \ell'_n)$  is a second gamete the diploid genotype is then denoted by  $GG'$ . The phenotypic value of the character in an individual whose genotype is  $GG'$  is computed additively as

$$x = \alpha_0 + \sum_{i=1}^n \alpha_i(\ell_i + \ell'_i) + e \quad (2)$$

where  $e$  is an environmental deviation, namely a random variable with null expectation and variance  $V_e$  that can depend on genotype.  $\alpha_0$  is a constant that represents some baseline phenotypic value, and can be of any sign, although, by definition  $\alpha_i > 0$  ( $i = 1, 2, \dots, n$ ).

Define

$$p_i = \mathcal{E}_p\{\ell_i\} \quad \text{and} \quad D_{ij} = \mathcal{E}_p[(\ell_i - p_i)(\ell_j - p_j)], \quad (3)$$

where  $\mathcal{E}_p$  is an expectation operator, so that  $p_i$  is the frequency of  $A_i$  and  $D_{ij}$  is the linkage disequilibrium between the  $i$ -th and  $j$ -th loci. Let  $\mathcal{E}_e$  denote the expectation with respect to the distribution of  $e$ . Under random union of gametes the mean of the phenotype defined by (2) is

$$\bar{x} = \alpha_0 + 2 \sum_{i=1}^n \alpha_i p_i \quad (4)$$

and its variance,  $V = \mathcal{E}_p \mathcal{E}_e[(x - \bar{x})^2]$ , satisfies

$$V = V_A + C_L + V_E \quad (5)$$

where

$$V_A = 2 \sum_i \alpha_i^2 p_i q_i \quad (6)$$

is an *additive variance*,

$$C_L = 2 \sum_i \sum_{j \neq i} \alpha_i \alpha_j D_{ij} \quad (7)$$

is an additional term due to linkage disequilibrium and  $V_E = \mathcal{E}_e\{V_e\}$ . The expansion (5) can be found in Weir et al. (1980) and the notation  $C_L$  in (7) is originally due to Bulmer (1980). A proof of the decomposition (5) is given in Appendix 1.

Selection occurs on the phenotype according to a quadratic stabilizing regime

$$w(x) = 1 - s(x - \theta)^2 \quad (8)$$

where  $\theta$  is an optimal phenotype at which fitness is maximized. The parameter  $s \geq 0$  is a measure of the intensity of the selection, the greater is  $s$  the stronger is the selection towards  $\theta$ . It is well known that the allele frequencies (3) under the specification (2) and (8), with constant environmental variance  $V_e$ , cannot achieve a stable polymorphic equilibrium (Wright, 1968; Lewontin, 1964). Generalization of the above framework may allow stable multilocus polymorphism under the selection regime (8). Here we consider the case where the environmental deviation depends on the genotype so that

$$V_e = E_0 - \sum_{i=1}^n \beta_i (\ell_i + \ell'_i - 2\ell_i \ell'_i), \quad (9)$$

where  $\beta_i$  are constants. This is a generalization of Lerner's model of homeostasis in which the greater is the heterozygosity the smaller is the variance. The scheme (9) differs from Lerner's by our inclusion of unequal contributions from different loci. The constant  $E_0$  is the variance corresponding to the state where all  $n$  loci are homozygous, and for  $V_e$  to be positive for every genotype we require  $E_0 > 2 \sum_{i=1}^n \beta_i$ . For future reference write  $V_\beta = 2 \sum_{i=1}^n \beta_i p_i q_i$ .

Substitution of (2) into (8) yields the fitness of the genotype  $GG'$

$$\begin{aligned} w(GG') = \mu + \sum_i [\tilde{K}_i (\ell_i + \ell'_i) + 2\tilde{L}_i \ell_i \ell'_i] \\ + \sum_{i \neq j} \sum_j \tilde{M}_{ij} (\ell_i + \ell'_i) (\ell_j + \ell'_j), \end{aligned} \quad (10)$$

where, as shown in Appendix 2,

$$\mu = 1 - s[(\alpha_0 - \theta)^2 + E_0]; \quad (11a)$$

$$\tilde{K}_i = s[\beta_i - \alpha_i^2 + 2(\theta - \alpha_0)\alpha_i] = sK_i, \quad (11b)$$

$$\tilde{L}_i = -s(\beta_i + \alpha_i^2) = sL_i, \quad \tilde{M}_{ij} = -s\alpha_i \alpha_j = sM_{ij} (i \neq j). \quad (11c)$$

The general dynamical system that determines the evolution of the gamete frequencies with fitnesses specified by (10) is extremely cumbersome to analyze. Using perturbation techniques, however, it is possible to determine the stationary states for small values of  $s$ . These have the form

$$p_i = p_i^0 + sp_i' + s^2 p_i'' + \dots \quad (12a)$$

$$D_{ij} = D_{ij}^0 + sD_{ij}' + s^2 D_{ij}'' + \dots \quad (12b)$$

Of course we seek solutions with  $0 < p_i^0 < 1$  ( $i = 1, 2, \dots, n$ ).

The properties of the perturbation in (12a, b) were demonstrated by Zhivotovsky and Gavrilets (1992, eqs (16c), (18), (19)). They showed that  $\{p_i^0\}$ ,  $\{p_i'\}$ ,  $\{D_{ij}^0\}$  and  $\{D_{ij}'\}$  satisfy

$$K_i + 2L_i p_i^0 + 4 \sum_{j \neq i} M_{ij} p_j^0 = 0 \quad (13)$$

$$D_{ij} = 0, \quad D_{ij}' = 2 \frac{M_{ij} p_i^0 p_j^0 q_i^0 q_j^0}{r_{ij}} \quad (14)$$

$$p_i^0 q_i^0 (2L_i p_i' + 4 \sum_{j \neq i} M_{ij} p_j') = -2(q_i^0 - p_i^0) \sum_{j \neq i} M_{ij} D_{ij}' \quad (15)$$

where  $r_{ij} \neq 0$  is the recombination fraction between loci  $i$  and  $j$ . We shall restrict our attention to terms in (12) of first order in  $s$ . In Appendix 3 we show that on substitution of (10) with (11) into (13), (14) and (15) the following versions of (12a) and (12b) are valid

$$p_i \approx \frac{1}{2} + (\theta - x_m) \frac{\alpha_i}{(\beta_i - \alpha_i^2)(1 + \varphi_1)} \left[ 1 - s \left( \tau_i - \frac{\varphi_1}{1 + \varphi_1} \hat{\tau} \right) \right] \quad (16)$$

$$D_{ij} = -2s \frac{\alpha_i \alpha_j p_i^0 q_i^0 p_j^0 q_j^0}{r_{ij}} (i \neq j). \quad (17)$$

Here  $x_m = \alpha_0 + \Sigma \alpha_i$  is the mean value of the quantitative character when  $p_i = 1/2$  (all  $i$ ). Thus  $\theta - x_m$  represents the deviation of the optimum value from the midpoint of the phenotypic range. The values  $\varphi_1, \tau_i, \hat{\tau}_i$  are positive values which are cumbersome to write down and are recorded in Appendix 3.

## Stability of Allele Frequencies

It was shown by Zhivotovsky and Gavrilets (1992) that for small values of  $\tilde{K}_i$ ,  $\tilde{L}_i$  and  $\tilde{M}_{ij}$  in the fitness expression (10) the polymorphic equilibrium specified by (16) and (17) is stable under the assumption  $\tilde{M}_{ij} = \gamma\alpha_i\alpha_j$  provided that  $\tilde{L}_i < 2\gamma\alpha_i^2$  ( $\gamma < 0$ ). Applying this result to our parameters (11) the condition for stability is

$$\beta_i > \alpha_i^2 \quad (i = 1, 2, \dots, n) \quad (18)$$

At the stable polymorphism there is linkage disequilibrium and, returning to (5) and (7) its contribution to the variance  $V$  is  $C_L$  with

$$C_L = -4s \sum_{i \neq j} \alpha_i^2 \alpha_j^2 p_i^0 q_i^0 p_j^0 q_j^0 / r_{ij} \quad (19)$$

## Discussion

The model defined by (2) and (8) with (9) is a possible alternative to existing models of quantitative variability. The stable polymorphic equilibrium specified by (16) and (17) has a number of interesting qualitative properties which we document below.

**Property 1.** At the stable state, irrespective of the values of the contributions of the different loci to the phenotypic value, the frequencies of alleles with a positive effect on a character under weak stabilizing selection are all on the same side of 1/2. That is, if  $\alpha_i > 0$  for all  $i$ , then all  $p_i$  are greater than 1/2 or all are less than 1/2. In fact, for sufficiently small selection

$$\text{if } \theta > x_M \text{ then every } p_i > 0.5$$

$$\text{if } \theta < x_m \text{ then every } p_i < 0.5.$$

This conclusion concerning the allele frequencies (which is a direct consequence of (16)) suggests that in a population subject to the constraints of the model there should be a preponderance of gametes (or chromosomes) with mainly  $A$ -alleles, or  $a$ -alleles, according to whether  $\theta > x_M$  or  $\theta < x_M$  respectively. Also from (16) we have

**Property 2.** At the polymorphic equilibrium the allele frequencies  $p_i$  are, in general, different from one another. Indeed from (16) we see that  $p_i \neq p_j$  for each  $s$  if  $\alpha_i / (\beta_i - \alpha_i^2) \neq$

$\alpha_j/(\beta_j - \alpha_j^2)$ . Thus, differences in the contributions from the different loci to the character under selection result in different allele frequencies. Even if all such contributions are equal, however, allele frequencies may differ one from the other if the loci are linked. Indeed, set  $\alpha_i = \alpha$ ,  $\beta_i = \beta$  and take  $n$  very large. Then in Appendix 4 we show that

$$p_i \approx \frac{1 + \theta_0}{2} - s \frac{\theta_0}{2} (V_A^0)^2 (V_\beta^0 - V_A^0)^{-1} \left( \frac{1}{\bar{r}} - \frac{1}{\tilde{r}_i} \right) \quad (20)$$

where  $\theta_0 = [(\theta - \alpha_0)/\alpha n] - 1$  is a relative measure of the deviation of the optimal value  $\theta$  from the midpoint  $x_m$ , with  $\theta_0 = 0$  if  $\theta = x_m$ ;  $\theta_0 = -1$  when  $\theta = \alpha_0$ , the minimum value of  $x$ ;  $\theta_0 = +1$  when  $\theta = \alpha_0 + 2n\alpha$ , the maximum value of  $x$ . Here

$$\tilde{r}_i = \left[ \frac{1}{n-1} \sum_{j \neq i} r_{ij}^{-1} \right]^{-1}, \quad \bar{r} = \left[ \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} r_{ij}^{-1} \right]^{-1} \quad (21)$$

are harmonic averages of the recombination fractions. Thus, in our model the  $p_i$ 's are all equal if and only if the loci contribute equally to the phenotype *and* all the loci recombine freely.

Property 1 has an obvious but interesting corollary. Suppose for example that  $\theta > x_m$  so that every  $p_i$  is larger than  $1/2$ . Then the majority of the gametes carry more  $A_i$ -alleles (+ alleles in Mather's terminology, i.e. alleles that increase the value of the trait) than  $a_i$ -alleles. Thus the proportions of complementary gametes are expected to be unequal.

**Property 3.** The greater the deviation of the optimal value,  $\theta$ , from the midpoint of the character range  $x_m$ , the less the frequency of the complementary gametes. In order to demonstrate this phenomenon we make the simplifying assumptions that all loci contribute equally to the trait, that allele frequencies are equal and that there is linkage equilibrium, that is terms order  $s$  are neglected from the allele frequencies. These assumptions do not qualitatively alter the validity of Property 3. Let  $p_i = p_*$ , for all  $i$  with  $0 < p_* < 1$ . Denote gametes carrying  $k$  alleles of type  $A_i$  by  $G(k)$ . Then the frequency  $P_k$  of  $G(k)$  is, neglecting terms  $O(\varepsilon_0)$ ,

$$P_k = \binom{n}{k} p_*^k q_*^{n-k}$$

where  $q_* = 1 - p_*$ .  $P_k$  achieves its maximum at  $k = np_* = k_*$ , say. Thus the relative frequencies of those gametes carrying close to  $k_*$  alleles of type  $A_i$  may be substantially

greater than those of their complementary gametes  $G(n - k_*)$ . To see this define  $\lambda_k = P_{n-k}/P_k$  as the ratio of the frequency of complementary gametes  $G(n - k)$  to  $G(k)$ . Then  $\lambda_k = \lambda_*^{2k-n}$  where  $\lambda_* = q_*/p_* < 1$ , and  $P_k$  is the probability that the ratio takes the value  $\lambda_k$  ( $k = 1, 2, \dots, n$ ). Fix a small number  $\eta$ . Then  $\lambda_k \leq \eta$  if  $(2k - n) \ln \lambda_* \leq \ln \eta$  i.e. if

$$k \geq \left( n + \frac{\ln \eta}{\ln \lambda_*} \right) / 2 = n_\eta,$$

say. Then the probability that  $\lambda_k \leq \eta$  is the probability that  $k \geq n_\eta$ , say  $P_\eta$ :

$$P_\eta = \sum_{k=n_\eta}^n P_k.$$

For example in the case  $p_* = 0.7$ ,  $n = 12$ ,  $\ln \lambda_* = -0.85$  so that if  $\eta = 0.1$  then  $n_\eta = 7.3$  and  $P_{0.1} > 0.72$ . If  $\eta = 0.2$ , then  $P_{0.2} = 0.88$ . Hence the ratio of the frequencies of complementary gametes is less than 0.1 (0.2) with probability 0.72 (0.88).

An alternative way to view the skewness of this distribution is to observe that for  $k = np_*$ , where  $P_k$  achieves its maximum, the value of  $\lambda_k$  is  $\lambda_*^{n(2p_*-1)} = \lambda_{\max}$ , say. For the same example  $p_* = 0.7$ ,  $n = 12$ ,  $\lambda_{\max} = 0.017$ , so that the modal expected value for the ratio of complementary gamete frequencies  $G(n - k)$  to  $G(k)$  is less than 2%.

Finally, observe that at the stationary state (16) and (17) entail that the disequilibrium is negative to first order in  $s$ . In fact, the form of  $C_L$  in (19) demonstrates

**Property 4.** At the stable polymorphic equilibrium both genotypic and phenotypic variance are less than those expected in the absence of linkage equilibrium. This follows because  $V_G = V_A + C_L$  and  $V = V_G + V_E$ . In (19)  $C_L$  is clearly negative.

The properties described above distinguish our treatment from earlier studies. Consider first the **polygene balance** model of Mather (1941, 1943, 1973), Mather and Jinks (1973). Although this was not expressed in terms of an evolutionary dynamical system, its main qualitative properties can be deduced. Implicit in Mather's discussion is the supposition of equal frequencies for complementary gametes, which differs from Property 3 above. Further, Mather supposes a genotypic structure of the type  $+ - + - \dots / - + - + \dots$  which entails negative linkage disequilibrium between loci 1 and 2 and between loci 2 and 3, but

positive linkage disequilibrium between loci 1 and 3 and 2 and 4 etc. Clearly this differs from Property 4 of our model which affirms the negative sign of the linkage disequilibrium.

The model of Bulmer (1974, 1976, 1985) is based on normal distribution theory and supposes linear regression equations that connect the genotypic values of relatives. The main qualitative property of Bulmer's treatment is that it results in negative linkage disequilibrium which reduces genotypic and phenotypic variance. This coincides with our Property 4. Bulmer's model, however, represents an unstable polymorphic equilibrium with equal contribution from all loci and equal allele frequencies of 0.5 (see Zhivotovsky and Gavrilets, 1991) in contrast to properties 1 and 2 of our model. Moreover, in the presence of linkage disequilibrium the distribution of the character under selection deviates from normality (Turelli and Barton, 1990) contrary to the assumption. It should also be noted that Bulmer's assumption that allele frequencies are all equal to 0.5 implies equal frequencies of complementary gametes, in contrast to Property 3 of our model.

Latter (1960), Kimura (1965), Lande (1975), Cavalli-Sforza and Feldman (1976), Felsenstein (1977), Fleming (1979), Turelli (1984), Barton (1986), Turelli and Barton (1990), Keightley and Hill (1990), Zhivotovsky and Gavrilets (1990) all examine versions of a model in which a stable equilibrium of allele frequencies is maintained by a balance between mutations that produce variation and stabilizing selection against it. This **mutation-selection balance** model allows a quantitative trait to maintain variability. These treatments, however, do not generally lead to predictions for the allele frequencies or relative importance of complementary gametes as in our properties 1, 2 and 3. The mutation-selection balance model also results in negligibly small linkage disequilibrium, which distinguishes it from both Bulmer's and our model (Property 4).

Our model is just one of a class that might result in stable polymorphism under stabilizing selection. For example, if the different alleles and/or loci were not additive in their contribution to the phenotype, stabilizing selection could, in some cases, result in polymorphism. We call this class of models **stabilizing selection-balance** models to emphasize that the variability does not demand mutation. We have demonstrated here that genetic polymorphism can be maintained under selection toward an optimum, and that the nature of the viability reflected on the quantitative character is compatible with

observations from natural and experimental populations.

## Appendix 1

Define the expectations  $\mathcal{E}_e$  and  $\mathcal{E}_p$  with respect to the distribution of  $e$  and  $\{\ell_i\}$  respectively. Then the variance  $V = \mathcal{E}_p \mathcal{E}_e (x - \bar{x})^2$

$$\begin{aligned} \mathcal{E}_p \mathcal{E}_e (x - \bar{x})^2 &= \mathcal{E}_p \mathcal{E}_e \left[ \sum_i \alpha_i (\ell_i + \ell'_i) + e - 2 \sum_i \alpha_i p_i \right]^2 \\ &= \mathcal{E}_p \mathcal{E}_e \left\{ \left[ \sum_i \alpha_i (\ell_i + \ell'_i) \right]^2 + e^2 + 4 \left[ \sum_i \alpha_i p_i \right]^2 \right. \\ &\quad \left. - 4 \left[ \sum_i \alpha_i (\ell_i + \ell'_i) \right] \left[ \sum_i \alpha_i p_i \right] \right\}. \end{aligned} \quad (A1)$$

since  $e$  has the null expectation. Hence, expanding,

$$\begin{aligned} \mathcal{E}_p \mathcal{E}_e (x - \bar{x})^2 &= \mathcal{E}_p \left\{ \sum_{i \neq j} \alpha_i \alpha_j (\ell_i + \ell'_i) (\ell_j + \ell'_j) \right\} + \mathcal{E}_p \left\{ \sum_i \alpha_i^2 (\ell_i + \ell'_i + 2\ell_i \ell'_i) \right\} \\ &\quad + \mathcal{E}_p \{V_e\} + 4 \left( \sum_i \alpha_i p_i \right)^2 - 4 \mathcal{E}_p \left\{ \left[ \sum_i \alpha_i (\ell_i + \ell'_i) \right] \left[ \sum_i \alpha_i p_i \right] \right\}. \end{aligned} \quad (A2)$$

The first term is

$$\begin{aligned} &\sum_{i \neq j} \alpha_i \alpha_j \{ \mathcal{E}_p \ell_i \ell_j + \mathcal{E}_p \ell_i \ell'_j + \mathcal{E}_p \ell'_i \ell_j + \mathcal{E}_p \ell'_i \ell'_j \} \\ &= \left[ \sum_{i \neq j} 2\alpha_i \alpha_j (D_{ij} + 2p_i p_j) \right], \\ &= \sum_{i \neq j} 2\alpha_i \alpha_j D_{ij} + 4 \left( \sum_i \alpha_i p_i \right)^2 - 4 \sum_i \alpha_i^2 p_i^2. \end{aligned} \quad (A3)$$

Incorporating (A3) into (A2) we obtain

$$\begin{aligned} V &= 2 \sum_{i \neq j} \alpha_i \alpha_j D_{ij} + 4 \left( \sum_i \alpha_i p_i \right)^2 - 4 \sum_i \alpha_i^2 p_i^2 \\ &\quad + 2 \sum_i \alpha_i^2 (p_i + p_i^2) - 4 \left( \sum_i \alpha_i p_i \right)^2 + V_e \\ &= 2 \sum_{i \neq j} \alpha_i \alpha_j D_{ij} + 2 \sum_i \alpha_i^2 p_i q_i + V_e \\ &= V_A + C_L + \mathcal{E}_p \{V_e\} \end{aligned} \quad (A4)$$

with  $V_A$  and  $C_L$  defined in (6) and (7).

## Appendix 2. Fitness with Lerner's Homeostasis

Combining (2), (8) and (9) we have

$$w(x) = 1 - s[\alpha_0 - \theta + \sum_i \alpha_i(\ell_i + \ell'_i) + e]^2 \quad (A5)$$

where  $e$  has zero expectation and variance

$$V_e = E_0 - \sum_i \beta_i(\ell_i + \ell'_i - 2\ell_i\ell'_i).$$

Expanding,

$$w(x) = 1 - s \left\{ (\alpha_0 - \theta)^2 + e^2 + 2e(\alpha_0 - \theta) + 2(\alpha_0 - \theta) \sum_i \alpha_i(\ell_i + \ell'_i) + 2e \sum_i \alpha_i(\ell_i + \ell'_i) + \left[ \sum_i \alpha_i(\ell_i + \ell'_i) \right]^2 \right\}.$$

Taking expectations with respect to the distribution of  $e$  we have

$$\begin{aligned} \bar{w}(x) &= 1 - s \left\{ (\alpha_0 - \theta)^2 + E_0 - \sum_i \beta_i(\ell_i + \ell'_i - 2\ell_i\ell'_i) + 2(\alpha_0 - \theta) \sum_i \alpha_i(\ell_i + \ell'_i) \right. \\ &\quad \left. + \left[ \sum_i \alpha_i(\ell_i + \ell'_i) \right]^2 \right\} \end{aligned} \quad (A6)$$

$$\begin{aligned} &= 1 - s[(\alpha_0 - \theta)^2 + E_0] + s \left\{ \sum_i [\beta_i - \alpha_i^2 + 2\alpha_i(\theta - \alpha_0)](\ell_i + \ell'_i) \right. \\ &\quad \left. - 2 \sum_i (\beta_i + \alpha_i^2)\ell_i\ell'_i - \sum_{i \neq j} \alpha_i\alpha_j(\ell_i + \ell'_i)(\ell_j + \ell'_j) \right\}. \end{aligned} \quad (A7)$$

which proves (10) with (11a, b, c).

Under the assumption of random union of gametes now take expectations with respect to the population frequency distribution, namely

$$\begin{aligned} \bar{w} &= 1 - s \left\{ (\alpha_0 - \theta)^2 + E_0 - 2 \sum_i \beta_i p_i q_i + 4(\alpha_0 - \theta) \sum_i \alpha_i p_i \right. \\ &\quad \left. + \mathcal{E}_p \left[ \sum_{i \neq j} \alpha_i \alpha_j (\ell_i + \ell'_i)(\ell_j + \ell'_j) + \sum_i \alpha_i^2 (\ell_i^2 + \ell_i'^2 + 2\ell_i \ell_i') \right] \right\} \end{aligned} \quad (A7)$$

$$\begin{aligned}
&= 1 - s \left\{ (\alpha_0 - \theta)^2 + E_0 - 2 \sum_i \beta_i p_i q_i + 4(\alpha_0 - \theta) \sum_i \alpha_i p_i \right. \\
&\quad \left. + \sum_{i \neq j} \sum \alpha_i \alpha_j \mathcal{E}_p(l_i l_j + l_i l'_j + l'_i l_j + l'_i l'_j) + 2 \sum_i \alpha_i^2 (p_i + p_i^2) \right\} \\
&= 1 - s \left\{ (\alpha_0 - \theta)^2 + E_0 - 2 \sum_i \beta_i p_i q_i + 4(\alpha_0 - \theta) \sum_i \alpha_i p_i \right. \\
&\quad \left. + 2 \sum_{i \neq j} \sum \alpha_i \alpha_j D_{ij} + 4 \left( \sum_i \alpha_i p_i \right)^2 - 4 \sum_i \alpha_i^2 p_i^2 + 2 \sum_i \alpha_i^2 (p_i + p_i^2) \right\} \\
&= 1 - s \left\{ (\alpha_0 - \theta)^2 + E_0 - 2 \sum_i \beta_i p_i q_i + 4(\alpha_0 - \theta) \sum_i \alpha_i p_i \right. \\
&\quad \left. + 2 \sum_{i \neq j} \sum \alpha_i \alpha_j D_{ij} + 4 \left( \sum_i \alpha_i p_i \right)^2 + 2 \sum_i \alpha_i^2 p_i q_i \right\}. \tag{A9}
\end{aligned}$$

On the other hand, Wright (1969 p. 106) expressed the mean fitness as

$$\bar{w} = 1 - s[(\bar{x} - \theta)^2 + V],$$

which entails, in our terminology,

$$\bar{w} = 1 - s[(\bar{x} - \theta)^2 + V_A + C_L + V_E] \tag{A10}$$

where  $\bar{x}$ ,  $V_A$  and  $C_L$  are defined by (4), (6) and (7) and  $V_E = E_0 - 2 \sum_i \beta_i p_i q_i$ .

### Appendix 3. Derivation of polymorphic equilibria (16), (17).

We use the relations (13), (14) and (15). (13) together with (11b) can be rewritten as

$$\begin{aligned}
 K_i &= \beta_i - \alpha_i^2 + 2(\theta - \alpha_0)\alpha_i = 2(\beta_i + \alpha_i^2)p_i^0 + 4 \sum_{j \neq i} \alpha_i \alpha_j p_j^0 \\
 &= 2(\beta_i - \alpha_i^2)p_i^0 + 4\alpha_i \sum_j \alpha_j p_j^0 \\
 &= 2(\beta_i - \alpha_i^2)p_i^0 + 4\alpha_i v^0
 \end{aligned}$$

where  $v^0 = \sum_j \alpha_j p_j^0$ . If  $\beta_i > \alpha_i^2$  (all  $i$ ) then  $K_i > 0$  and

$$p_i^0 = \frac{K_i}{2(\beta_i - \alpha_i^2)} - \frac{2\alpha_i v^0}{(\beta_i - \alpha_i^2)}. \quad (A11)$$

Multiply both sides of (A11) by  $\alpha_i$  and sum over  $i$ . Then

$$v^0 = \sum_i \alpha_i K_i / 2(\beta_i - \alpha_i^2) - 2[\sum \alpha_i^2 / (\beta_i - \alpha_i^2)] v_0,$$

and

$$\begin{aligned}
 v_0 &= \frac{\sum_i \alpha_i K_i / 2(\beta_i - \alpha_i^2)}{1 + 2 \sum_i \alpha_i^2 / (\beta_i - \alpha_i^2)} \\
 &= A / 2(1 + \varphi_1),
 \end{aligned} \quad (A12)$$

where

$$\begin{aligned}
 A &= \sum_i \frac{\alpha_i [\beta_i - \alpha_i^2 + 2(\theta - \alpha_0)\alpha_i]}{\beta_i - \alpha_i^2} \\
 &= \sum_i \alpha_i + (\theta - \alpha_0)\varphi_1
 \end{aligned} \quad (A13)$$

with

$$\varphi_1 = 2 \sum_i \alpha_i^2 / (\beta_i - \alpha_i^2). \quad (A14)$$

Returning to (A11) we have therefore

$$\begin{aligned}
 p_i^0 &= \frac{1}{2} + \frac{(\theta - \alpha_0)\alpha_i}{\beta_i - \alpha_i^2} - \frac{\alpha_i}{\beta_i - \alpha_i^2} \frac{A}{(1 + \varphi_1)} \\
 &= \frac{1}{2} + \frac{B_i}{\beta_i - \alpha_i^2},
 \end{aligned} \quad (A15)$$

where  $B_i = \alpha_i[\theta - \alpha_0 - A/(1 + \varphi_1)] = \alpha_i B_0$ , say.

Using (11c) write (14) as

$$D'_{ij} = -2\alpha_i \alpha_j p_i^0 q_i^0 p_j^0 q_j^0 / r_{ij}. \quad (\text{A16})$$

Then, returning to (15) with (11c) we have

$$p_i^0 q_i^0 [-(\beta_i + \alpha_i^2)p'_i - 2 \sum_{j \neq i} \alpha_i \alpha_j p'_j] = (q_i^0 - p_i^0) \sum_{j \neq i} \alpha_i \alpha_j D'_{ij} \quad (\text{A17})$$

which can be rewritten, using (A16) and  $T' = \sum_j \alpha_j p'_j$ , as

$$[(\beta_i - \alpha_i^2)p'_i + 2\alpha_i T'] = (q_i^0 - p_i^0) \alpha_i^2 2 \sum_{j \neq i} \frac{\alpha_j^2 p_j^0 q_j^0}{r_{ij}},$$

since  $p_i^0 q_i^0 \neq 0$ . Hence, since  $q_i^0 - p_i^0 = -2B_i/(\beta_i - \alpha_i^2)$ ,

$$p'_i = \frac{-2\alpha_i T'}{\beta_i - \alpha_i^2} - \frac{2B_i}{(\beta_i - \alpha_i^2)^2} \alpha_i^2 T_i \quad (\text{A18})$$

where  $T_i = 2 \sum_{j \neq i} \alpha_j^2 p_j^0 q_j^0 / r_{ij}$ . Now multiply (A18) by  $\alpha_i$  and sum to produce

$$T' = N/(1 + \varphi_1) \quad (\text{A19})$$

where  $N = -2 \sum_i B_i \alpha_i^3 T_i / (\beta_i - \alpha_i^2)^2$ . Thus the complete solution (A18) may be expanded as

$$p'_i = \frac{4\alpha_i}{\beta_i - \alpha_i^2} (1 + \varphi_1)^{-1} \sum_{j=1}^n \frac{B_j \alpha_j^3 T_j}{(\beta_j - \alpha_j^2)^2} - \frac{2\alpha_i^2 B_i T_i}{(\beta_i - \alpha_i^2)^2}. \quad (\text{A20})$$

Define  $\tau_i = \alpha_i^2 T_i / (\beta_i - \alpha_i^2)$ . Then the value

$$\bar{\tau} = \frac{1}{\varphi_1} 2 \sum_j \frac{\alpha_j^2}{\beta_j - \alpha_j^2} \tau_j. \quad (\text{A21})$$

is an average of these  $\tau$ 's with weights  $2\alpha_j^2/\varphi_1(\beta_j - \alpha_j^2)$  whose sum is unity. Upon substitution of  $\tau_i$  and  $\bar{\tau}_i$  into (A18), using (A19) we find

$$\begin{aligned} p'_i &= \frac{B_0 \alpha_i \varphi_1}{(1 + \varphi_1)(\beta_i - \alpha_i^2)} \left[ (\bar{\tau} - \tau_i) - \frac{1}{\varphi_1} \tau_i \right] \\ &= \frac{B_0 \alpha_i}{(1 + \varphi_1)(\beta_i - \alpha_i^2)} [\varphi_1 (\bar{\tau} - \tau_i) - \tau_i]. \end{aligned} \quad (\text{A22})$$

Since  $A = \sum_i \alpha_i + (\theta - \alpha_0)\varphi_1$  and  $B_0 = \theta - \alpha_0 - A/(1 + \varphi_1)$  we have

$$B_0 = \frac{\theta - x_m}{1 + \varphi_1} \quad (\text{A22})$$

where  $x_m = \alpha_0 + \sum_i \alpha_i$ . Hence

$$\begin{aligned} p'_i &= \frac{(\theta - x_m)}{1 + \varphi_1} \frac{\alpha_i}{\beta_i - \alpha_i^2} \frac{[\varphi_1(\tilde{\tau} - \tau_i) - \tau_i]}{1 + \varphi_1} \\ &= -\frac{\theta - x_m}{1 + \varphi_1} \frac{\alpha_i}{\beta_i - \alpha_i^2} \left[ \tau_i - \frac{\varphi_1 \tilde{\tau}}{1 + \varphi_1} \right]. \end{aligned} \quad (\text{A23})$$

Now combine (A15) and (A23) with (12a) to yield (16).

#### Appendix 4. Derivation of (20) from (16)

Suppose that  $\alpha_i = \alpha$ ,  $\beta_i = \beta$ . Then  $\varphi_1$  reduces to  $2n\alpha^2/(\beta - \alpha^2)$  so that as  $n \rightarrow \infty$ ,  $\varphi_1/(1 + \varphi_1) \rightarrow 1$ . For  $n$  large we therefore take  $1 + \varphi_1 \approx \varphi_1$ . Now with these assumptions the weights of  $\tau_i$  used to compute  $\tilde{\tau}$  in (A21) reduce to

$$\frac{2\alpha^2}{\varphi_1(\beta - \alpha^2)} = \frac{1}{n},$$

so that  $\tilde{\tau} = \sum \tau_i/n$ . Write

$$R_i = \frac{\alpha_i}{\beta_i - \alpha_i^2} \frac{1}{1 + \varphi_1} \left( \tau_i - \frac{\varphi_1}{1 + \varphi_1} \tilde{\tau} \right)$$

(see eqn (16))

$$= (\tau_i - \tilde{\tau})/2\alpha n$$

under the assumption made above. But  $\tau_i = \alpha^2 T_i / (\beta - \alpha^2)$ , where, since  $p_j^0 = p^0$

$$T_i = 2 \sum_{j \neq i} \alpha^2 p_j^0 q_j^0 / \tau_{ij} = 2(n-1)\alpha^2 p^0 q^0 \tilde{r}_i^{-1}.$$

That is

$$\tau_i = \frac{\alpha^2}{\beta - \alpha^2} 2(n-1)\alpha^2 p^0 q^0 \tilde{r}_i^{-1} \approx (2n\alpha^2 p^0 q^0) \frac{\alpha^2}{\beta - \alpha^2} \tilde{r}_i^{-1}$$

and

$$\tilde{\tau} = 2n\alpha^2 p^0 q^0 \frac{\alpha^2}{\beta - \alpha^2} \tilde{r}^{-1}.$$

Finally

$$R_i = \frac{1}{2\alpha n} \frac{\alpha^2}{\beta - \alpha^2} 2n\alpha^2 p^0 q^0 (\tilde{r}_i^{-1} - \tilde{r}^{-1}).$$

Define  $\theta_0$  according to the relation

$$\theta - x_m = \theta - \alpha_0 - n\alpha \stackrel{\text{def}}{=} \alpha n \theta_0.$$

Then

$$-s(\theta - x_m)R_i = -s\theta_0 \frac{\alpha^2}{2(\beta - \alpha^2)} V_A^0 (\tilde{r}_i^{-1} - \tilde{r}^{-1})$$

where  $V_A^0 = 2n\alpha^2 p^0 q^0$  and  $V_\beta^0 \stackrel{\text{def}}{=} 2n\beta p^0 q^0$ . Hence

$$\begin{aligned}
-s(\theta - x_m)R_i &= -\frac{s\theta_0}{2} \frac{2n\alpha^2 p_0 q_0}{2n\beta p_0 q_0 - 2n\alpha^2 p_0 q_0} V_A^0 (\tilde{r}_i^{-1} - \tilde{r}^{-1}) \\
&= -\frac{s\theta_0}{2} \frac{(V_A^0)^2}{V_\beta^0 - V_A^0} (\tilde{r}_i^{-1} - \tilde{r}^{-1}).
\end{aligned}$$

Finally

$$(\theta - x_m) \frac{\alpha}{(\beta - \alpha^2)(1 + \varphi_1)} = \alpha n \theta_0 \frac{\alpha}{\beta - \alpha^2} \frac{\beta - \alpha^2}{2\alpha^2 n} = \frac{\theta_0}{2}$$

so that

$$p_i = \frac{1 + \theta_0}{2} - \frac{s\theta_0}{2} \frac{(V_A^0)^2}{V_\beta^0 - V_A^0} (\tilde{r}_i^{-1} - \tilde{r}^{-1})$$

as claimed in (20).

## References

- Barton, N.H. 1986. The effect of linkage and density-dependent regulation on gene flow. *Hereditary* **57**: 415-426.
- Barton, N.H., M. Turelli. 1990. Evolutionary quantitative genetics: how little do we know? *Ann. Rev. Genet.* **23**:
- Bulmer, M.G. 1972. The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genet. Res.* **19**: 17-25.
- Bulmer, M.G. 1974. Linkage disequilibrium and genetic variability. *Genet. Res.* **23** 281-289.
- Bulmer, M.G. 1976. The effect of selection on genetic variability: a simulation study. *Genet. Res.* **28**: 101-117.
- Bulmer, M.G. 1980. *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford.
- Burger, R. 1988. Mutation-selection balance and continuum of  $n$ -alleles models. *Math. Biosci.* **91**: 67-83.
- Cavalli-Sforza, L.L., and M.W. Feldman. 1976. Evolution of continuous variation: direct approach through joint distribution of genotypes and phenotypes. *Proc. Natl. Acad. Sci.* **73**: 1689-1692.
- Cockerham, C.C. 1954. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859-882.
- Crow, J.F., and M. Kimura. 1970. *An Introduction to Population Genetics Theory*. New York, Harper Row.
- Dobzhansky, Th. 1970. *Genetics of Evolutionary Process*. New York: Columbia University Press.
- Feldman, M.W., I. Franklin, and G.J. Thomson. 1974. Selection in complex genetic systems I. The symmetric equilibria of the three-locus symmetric viability model. *Genetics* **76**: 135-162.
- Felsenstein, J. 1977. Multivariate normal genetic models with a finite number of loci. p. 227-245. In Pollack, E., O. Kempthorne, and T.B. Bailey Jr. (eds.) *Proc. Int. Conference*

- on Quantitative Genetics*. Iowa State University Press, Ames, Iowa.
- Fisher, R.A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh* 52: 399-433.
- Fleming, W.H. 1979. Equilibrium distributions of continuous polygenic traits. *SIAM J. Appl. Math.* 36: 148-168.
- Futuyma, D.A. 1986. *Evolutionary Biology*. Sinauer Associates Inc., Sunderland, Mass.
- Hastings, A. 1986. Multilocus population genetics with weak epistasis. II. Equilibrium properties of multilocus selection: what is the unit of selection? *Genetics* 112: 157-171.
- Karlin, S. 1979. Principles of polymorphism and epistasis for multilocus systems. *Proc. Natl. Acad. Sci. USA* 76: 541-545.
- Karlin, S. and M.W. Feldman. 1970. Linkage and selection: two-locus symmetric viability model. *Theor. Pop. Biol.* 1: 39-71.
- Keightley, P.D., and W.G. Hill. 1990. Variation maintained in quantitative traits with mutation-selection balance: pleiotropic side-effect on fitness trait. *Proc. Royal Soc., London B* 242: 95-100.
- Kempthorne, O. 1957. *An Introduction to Genetic Statistics*. John Wiley, Chapman & Hall, London, New York.
- Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. USA* 54: 731-736.
- Lande, R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* 26: 221-235.
- Lerner, I.M. 1954. *Genetic Homeostasis*. Wiley, New York.
- Lewontin, R.C. and K. Kojima. 1960. The evolutionary dynamics of complex polymorphisms. *Evolution* 14: 458-472.
- Lewontin, R.C. 1964. The interaction of selection and linkage. II. Optimum models. *Genetics* 50: 757-782.
- Lewontin, R. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- Mather, K. 1973. *Genetical Structure of Populations*. Chapman and Hall, London.

- Mather, K. 1942. The balance of polygene combinations. *J. Gent.* 43: 309-336.
- Mather, K. and J.L. Jinks. 1982. *Biometrical Genetics*. Cambridge University Press, Cambridge.
- Nagylaki, T. 1976. The evolution of one- and two-locus systems. *Genetics* 83: 583-600.
- Nagylaki, T. 1989. The maintenance of genetic variability in two-locus models of stabilizing selection. *Genetics* 122: 235-248.
- Schmalhausen, I.I. 1986. *Factors of Evolution. The Theory of Stabilizing Selection*. The University of Chicago Press, Chicago, London.
- Slatkin, M. 1972. On treating the chromosome as the unit of selection. *Genetics* 72: 157-168.
- Thoday, J.M. 1961. Location of polygenes. *Nature* 191: 368-370.
- Thompson, J.N. Jr., J.J. Hellack, and R.R. Tucker. 1991. Evidence for balanced linkage of X chromosome polygenes in a natural population of *Drosophila*. *Genetics* 127: 117-123.
- Turelli, M. 1984. heritable genetic variation via mutation-selection balance: Lersch's zeta meets the abdominal bristle. *Theor. Popul. Biol.* 25: 138-193.
- Turelli, M. 1988. Population genetic models for polygenic variation and evolution. p. 601-618. In Weir, B.S., E.J. Eisen, M.M. Goodman, and G. Namkoong (eds.) *Proc. of the Second Intern. Conf. on Quantitative Genetics*. Sinauer, Sunderland, Mass.
- Turelli, M., and Barton, N.H. 1990. Dynamics of polygenic characters under selection. *Theor. Popul. Biol.* 38: 1-57.
- Weir, B.S., C.C. Cockerham, and J. Reynolds. 1980. The effects of linkage and linkage disequilibrium on the covariances of noninbred relatives. *Heredity* 45: 351-359.
- Wright, S. 1968. *Evolution and Genetics of Populations*. Vol. 2. Univ. of Chicago Press, Chicago.
- Zhivotovsky, L.A., and S. Yu. Gavrilets. 1990. Stabilizing selection and linkage disequilibrium. *Soviet Genetics* 26: N2 222-231 (English translation).
- Zhivotovsky, L., and S. Gavrilets. 1992. Quantitative variability and multilocus polymorphism under epistatic selection. *Theor. Pop. Biol.* (To appear)

