Genet. Res., Camb. (1987), **49**, *pp.* 157–173 *With 2 text-figures Printed in Great Britain*

Adaptive landscapes, genetic distance and the evolution of quantitative characters

N. H. BARTON
Department of Genetics and Biometry, University College, 4 Stephenson Way, London NW1 2HE, U.K.

MICHAEL TURELLI
Department of Genetics, University of California, Davis, California 95616, U.S.A.

(*Received 6 August 1986 and in revised form 21 November 1986*)

Summary

The maintenance of polygenic variability by a balance between mutation and stabilizing selection has been analysed using two approximations: the 'Gaussian' and the 'house of cards'. These lead to qualitatively different relationships between the equilibrium genetic variance and the parameters describing selection and mutation. Here we generalize these approximations to describe the dynamics of genetic means and variances under arbitrary patterns of selection and mutation. We incorporate genetic drift into the same mathematical framework.

The effects of frequency-independent selection and genetic drift can be determined from the gradient of log mean fitness and a covariance matrix that depends on genotype frequencies. These equations describe an 'adaptive landscape', with a natural metric of genetic distance set by the covariance matrix. From this representation we can change coordinates to derive equations describing the dynamics of an additive polygenic character in terms of the moments (means, variances, ...) of allelic effects at individual loci. Only under certain simplifying conditions, such as those derived from the Gaussian and house-of-cards approximations, do these general recursions lead to tractable equations for the first few phenotypic moments. The alternative approximations differ in the constraints they impose on the distributions of allelic effects at individual loci. The Gaussian-based prediction that evolution of the phenotypic mean does not change the genetic variance is shown to be a consequence of the assumption that the allelic distributions are never skewed. We present both analytical and numerical results delimiting the parameter values consistent with our approximations.

1. Introduction

Many of the characters important in adaptation and artificial selection are continuously distributed. Their phenotypic distributions are often approximately Gaussian, or may become so when measured on an appropriate scale. Beyond experiments suggesting polygenic inheritance, the genetic basis of variation of such traits is almost always unknown. Given our ignorance, it would be useful to find conditions under which the evolution of these traits can be predicted from measurable variables, such as the phenotypic mean and the genetic variance, in a way that is independent of genetic details. The general question is: when can we ignore the genetics underlying heritable variation in quantitative traits?

In the short term, artificial selection experiments suggest that the additive genetic variance may remain roughly constant, so that the mean will change at a rate proportional to the selection pressure on the mean. Over longer time-scales, the Gaussian phenotypic analyses popularized by Lande (1976, 1979) provide a valuable description of the evolutionary dynamics of character means, provided that one is willing to assume that the genetic variances and covariances are known. However, this approach cannot predict the dynamics of the variance. These dynamics are fundamental to long-term selection response and to understanding the processes that maintain polygenic variation. We present a general description of the dynamics of the mean, variance and higher-order moments of a polygenic trait. This description provides the conditions under which phenotypic dynamics may be predicted without complete knowledge of the underlying genetics.

Crow & Kimura (1964) introduced a model in which variation in a continuous character is maintained by a balance between mutation and stabilizing selection. They assumed an infinite number of segregating alleles at each locus, giving a continuous distribution of allelic effects. Kimura (1965) showed that the equilibrium distribution is approximately Gaussian, under the implicit assumption that the effects of new mutations are small relative to the existing variation at the locus (Turelli, 1984). Lande (1975, 1980) used this result to analyse the effects of linkage and pleiotropy. By *assuming* that the effects of all the loci follow a multivariate Gaussian distribution, the evolution of a polygenic trait can be described completely in terms of the mean vector and covariance matrix of this distribution. This greatly simplifies genetic analyses of many important evolutionary problems ranging from sexual selection to character displacement (e.g. Lande, 1977, 1981; Felsenstein, 1979; Kirkpatrick, 1982, 1985; Slatkin, 1980).

Turelli (1984, 1985, 1986) has criticized this Gaussian model on two grounds: first, per-locus mutation rates are unlikely to be high enough to maintain a Gaussian distribution of allelic effects at individual loci, and secondly, there are unlikely to be enough alleles at each locus to support the continuum-of-alleles approximation for more than a single character. Turelli (1984) developed an alternative approximation for the equilibrium of Crow & Kimura's continuum-of-alleles model. His approximation is based on the empirically motivated assumption that the effects of new mutations at a locus are generally much greater than the existing genetic variance at the locus. In a different context, Kingman (1978) described such mutations as 'bring(ing) down the evolutionary "house of cards"', and Turelli (1984) showed that, when per-locus mutation rates are low relative to selection intensity, a 'house of cards' approximation accurately describes the equilibrium of allelic effects under the model analysed by Kimura (1965) and Lande (1975). Despite its derivation from a continuum-of-alleles model, this approximation produces predictions for the equilibrium genetic variance that are essentially identical to those from models in which only a few alleles segregate at each locus (e.g. Wright, 1935; Latter, 1960; Bulmer, 1972; Turelli, 1984; Slatkin, 1987), provided the alleles responsible for polygenic variation are usually rare.

The assumption that loci are near fixation is not obviously identical to the assumption that new mutations swamp the existing variability at a locus. We will therefore distinguish them in this paper, by referring to the former as the 'rare alleles' approximation, and the latter as the 'house of cards' approximation. In models with a continuum of alleles, each allele is infinitesimally rare. The assumption that alleles producing variance are rare is then replaced by the more general assumption that the distribution of allelic effects is highly leptokurtic: i.e. that the fourth central moment is much larger than the square of the variance.

Here we extend the Gaussian and house-of-cards approximations to describe the response of the mean and the variance to arbitrary selection pressures, and to sampling drift. We also consider the genetic consequences of non-Gaussian distributions of phenotypes. Our general approach to modelling selection is described in the next section. It combines Wright's notion of an 'adaptive landscape' with a natural measure of 'genetic distance' (Akin, 1979). It leads to equations for the changes of the population mean, variance and skew that depend on the relation between the mean fitness of the population and those moments. Expressing the evolution of quantitative characters in terms of such an adaptive landscape aids calculation of (for example) the equilibrium distribution of genotypes and phenotypes under drift and the probability of shifts between alternative adaptive peaks. However, even in the absence of dominance, epistasis and genetic drift, the conditions under which a simple formulation is possible are biologically restrictive. Thus it may be that no more than a crude approximation of the genetic dynamics underlying phenotypic evolution is generally possible.

2. General approach

A random mating population can be described by the frequencies of all possible gametes. We will simplify by assuming that recombination rates (r) are large enough relative to selection intensities (s) and effective population sizes (N) so that the effects of linkage disequilibrium are negligible (i.e. s, $1/N \ll r$). The population can then be described by the frequencies of alleles at each locus. Numerical results that support this approximation for reasonable parameter values appear in Bulmer (1972), Turelli (1984) and Lynch & Hill (1986); see also Wright (1965). We are concerned here with an additive polygenic character, and so it will be convenient to replace the set of allele frequencies at a locus by the distribution of effects of the alleles on the character. The distribution will be continuous under the Crow & Kimura (1964) model.

We seek to describe the population in terms of the moments of the phenotypic distribution, rather than by the complete set of allele frequencies. This may be achieved in three steps. First, the coordinates are transformed from the allele frequencies to the mean, variance and higher moments of the distributions of allelic effects at each locus. This step is exact, because it only involves a change in the way we label different states of the population. For example, Fig. 1 shows three alternative representations of a population in which three alleles with effects $(-1, 0, 1)$ are segregating at a single locus. The population can be represented by the frequencies of all three alleles (Fig. 1*a*), by two of the three allele frequencies (Fig. 1*b*), or by the mean and variance of diploid individuals (Fig.

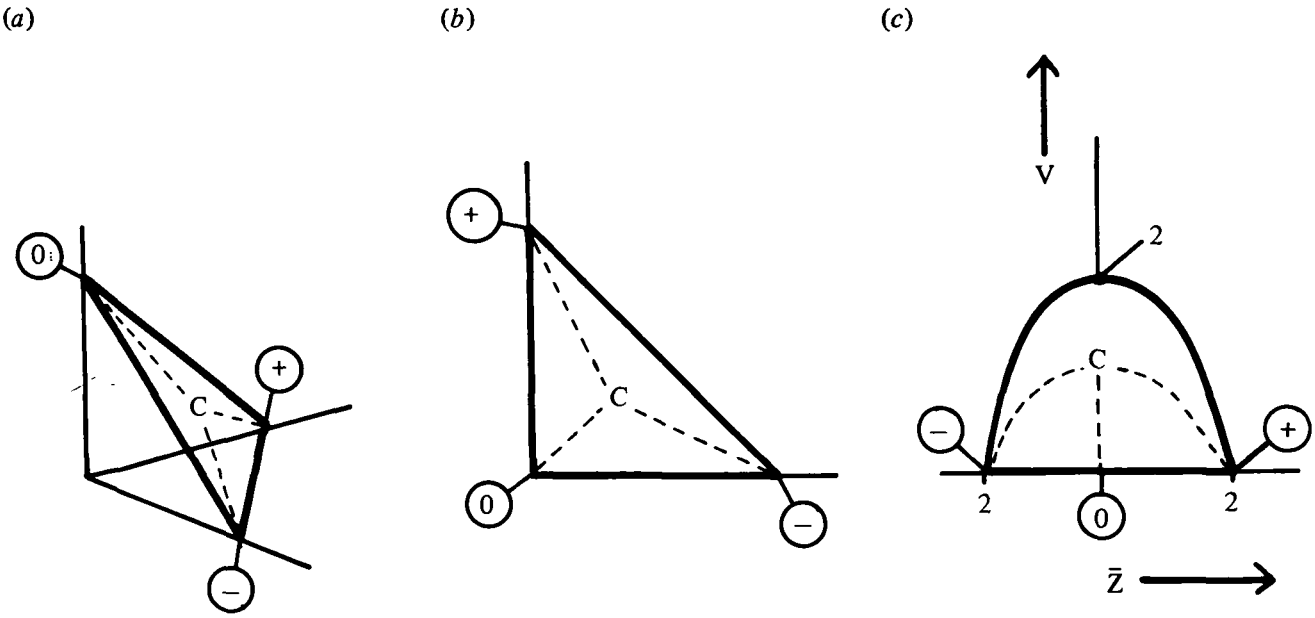

Fig. 1. The state of a population that segregates at a single locus for three alleles with effects $\{-1, 0, 1\}$ can be represented in many ways. (*a*) Graphs the frequencies of the three alleles; the population is confined to the two-dimensional plane on which allele frequencies sum to one. (*b*) Graphs the frequencies of two of the alleles, and (*c*) graphs the mean and variance of allelic effects. The three points −, 0, + (enclosed in circles) correspond to fixation of the three alleles. The point C corresponds to allele frequencies (1/3, 1/3, 1/3).

1*c*). The transformation to moments is simplified by assuming that selection is sufficiently weak that it allows a continuous-time approximation. The second step is to combine the recursions for individual loci to produce recursions for the phenotypic moments. Finally, conditions are found under which the dynamics of the first two or three moments are independent of the higher moments and the moments of effects at individual loci.

The transformation from allele frequencies to moments is facilitated by writing the equations for selection and drift in terms of an 'adaptive landscape' (Wright, 1935). In our notation the rate of change of the population due to selection is the product of a matrix and the gradient of log mean fitness. The covariance caused by sampling drift in each generation is given by the same matrix, divided by the effective population size. When selection is weak enough, and population size large enough, for the diffusion approximation to be accurate, these two equations suffice to describe the stochastic evolution of the population. The advantage of using Wright's formulation of selection and drift is that the equations take the same form, regardless of the coordinates being used: technically, they are 'covariant' (Adler, Bazin & Schiffer, 1975). Whilst Wright's equations have been used to model allele frequencies at a single locus, or at a set of independently segregating loci, it is not widely appreciated that they also apply to the effects of selection and sampling drift on gamete frequencies. Hence they also give a compact description of the effects of selection and drift on any transformation of gamete frequencies: for example, of allele frequencies and higher-order linkage disequilibria (Akin, 1979; Barton, 1987). We use the adaptive landscape here in a similar way, to give a compact description of the effects of selection and drift on the moments of an additive polygenic character; we approximate by neglecting linkage disequilibria, so as to avoid the complex effects of recombination. Our approach can be seen as an extension of Lande's (1976) analysis of phenotypic evolution to include changes in moments other than the mean.

3. Adaptive landscapes

If recombination rates are large enough relative to selection intensities, the response to selection of any population whose genotypes have constant relative fitnesses is proportional to the gradient in log mean fitness (Wright, 1935, 1969, chs. 3, 4; Akin, 1979, p. 51). Let $x = (x_1, x_2, \ldots)$ denote an arbitrary vector of variables that suffice to describe the evolution of a population without linkage disequilibrium; these might be allele frequencies or moments of distributions of allelic effects. If the population is large and diploid, mating is random, and selection is weak enough to permit a continuous-time approximation of the standard discrete-generation difference equations, we will show that there is a positive definite matrix $G = (g_{ij})$ such that

$$dx_i/dt \simeq \Sigma_j g_{ij} \partial(\ln \overline{W})/\partial x_j, \qquad (3.1)$$

$\overline{W}$ denoting the mean fitness of the population.

In general, the matrix G will vary with x. It is easiest to interpret when x is a vector of allele frequencies at

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

a single locus. For each allele at this locus, allele i say, define a genotypic indicator variable, denoted l_i, that takes the value 1 for allele i homozygotes, $\frac{1}{2}$ for allele i heterozygotes, and 0 for all other genotypes (cf. Price, 1970). (An alternative procedure, which would lead to essentially the same results, assigns haploid gametes carrying allele i a value $l_i = 1$, and all other gametes a value 0.) With this definition, the allele frequency p_i is just the expectation of l_i taken with respect to the genotype frequencies in the population. Let g_{ij} denote the covariance between l_i and l_j. Taking expectations with respect to the genotype frequencies in an infinite, random-mating, diploid population, we find

$$g_{ij} = p_i(\delta_{ij} - p_j)/2. \tag{3.2}$$

Here δ_{ij} is the Kronecker delta, equal to 1 if $i = j$, and 0 otherwise. We will refer to this G as the 'genetic indicator covariance matrix' or simply the 'covariance matrix'. Its role in (3.1) is precisely that played by the usual additive genetic covariance matrix in the equations for the dynamics of the means of multiple characters under selection. In fact the equations are identical, the characters here being the allelic indicators (l_i).

The standard equations for selection on a locus with s alleles are

$$\Delta p_i = p_{i,t+1} - p_{i,t} = p_i(\overline{W}_i - \overline{W})/\overline{W} \quad \text{for} \quad i = 1, \ldots, s, \tag{3.3}$$

with $\overline{W}_i = \Sigma_j p_j W_{ij}$, $\overline{W} = \Sigma_i p_i \overline{W}_i$, and W_{ij} the fitness of genotype ij. A straightforward calculation shows that these are equivalent to

$$\Delta p_i = \Sigma_j g_{ij} \partial(\ln \overline{W})/\partial p_j, \tag{3.4}$$

where ln denotes the natural logarithm. Thus for one-locus selection the matrix G in equation 3.1 is the variance–covariance matrix for the genotypic variables l_i, and equation 3.1 simply involves approximating Δp_i by dp_i/dt. (Alternative interpretations of the derivative $\partial(\ln \overline{W})/\partial p_j$ are discussed in Appendix A.)

The role of our assumption of linkage equilibrium becomes clear if we consider allele frequencies under multilocus selection. As shown by Ewens & Thomson (1977), equation 3.3 still applies if the fitnesses W_{ij} are replaced by 'induced fitnesses', denoted W^*_{ij}. An induced one-locus fitness is calculated by averaging the fitnesses of all the multilocus genotypes that contain the specified one-locus genotype. In the average, each fitness is weighted by the frequency of the associated multilocus genotype among individuals with the specified one-locus genotype. Linkage equilibrium simplifies matters because it makes the induced fitnesses independent of the allele frequencies at the locus being considered. Thus the calculations leading to equation 3.4 are unchanged. Linkage equilibrium also implies that the indicator variables for distinct loci are independent, so that if i and j indicate alleles at distinct loci, $g_{ij} = 0$. Hence equation 3.4 describes allele frequency dynamics under multilocus selection, with the summation extending over alleles at all loci, but with non-zero contributions coming only from one locus (see Wright, 1969, ch. 4).

Our continuous-time approximation simplifies the analysis of the more general case in which the x_i denote functions of the allele frequencies. A standard change of variables shows that equation 3.1 still applies with

$$g(x)_{ij} = \Sigma_{\alpha\beta}(\partial x_i/\partial p_\alpha)\, g(p)_{\alpha\beta}(\partial x_j/\partial p_\beta), \quad \text{or} \tag{3.5a}$$

$$G(x) = AG(p)\,A^T, \tag{3.5b}$$

where $A = (\partial x_i/\partial p_j)$ is the Jacobian matrix of the transformation from p to x and A^T denotes its transpose. Because the covariance matrix is necessarily positive definite,

$$d(\ln \overline{W})/dt = \Sigma_{ij}[\partial(\ln \overline{W})/\partial x_i]\, g_{ij}[\partial(\ln \overline{W})/\partial x_j] \geqslant 0. \tag{3.6}$$

Thus the population must evolve under selection in such a way that the mean fitness never decreases. This *Fundamental Theorem of Natural Selection* (Fisher, 1930; Akin, 1979, p. 70) implies that stable equilibria are at local peaks in mean fitness.

The representation of selection in terms of the gradient in log mean fitness (equation 3.1) is also useful when sampling drift is taken into account. From the usual Wright–Fisher model for genetic drift in a population of effective size N, we obtain

$$\operatorname{cov}(\delta p_i, \delta p_j) = g_{ij}/N. \tag{3.7}$$

Thus the covariance between the allele frequency fluctuations that are produced by sampling drift each generation is proportional to the covariance, g_{ij}. When i and j refer to alleles at the same locus, equations 3.2 and 3.7 give the standard covariance expression from multiple-allele diffusion approximations of genetic drift (cf. Watterson, 1977). For alleles at different loci, the approximation $\operatorname{cov}(\delta p_i, \delta p_j) = 0$ depends on our assumption of linkage equilibrium. As before, a simple coordinate transformation leads directly to the covariance caused by drift for an arbitrary set of characters, x_i.

Because the same matrix describes the effects of drift (equation 3.7) and the response to selection (equation 3.1), analysis of the stochastic behaviour of populations is greatly simplified. Any system whose dynamics can be represented by equations with the same form as those above has an asymptotic probability density given by Wright's (1937, 1969, chs. 13, 14) formula

$$\Psi(p) = C\overline{W}^{2N}/\det(G) \tag{3.8a}$$

(where C is a normalization constant). Without reversible mutation, all genetic variation must eventually be lost. Then, equation 3.8*a* is dominated by singularities at the fixation states: it describes the asymptotic frequency distribution for unfixed loci (see Ewens, 1964). If the mutation rate from allele i to allele

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

j, μ_{ij}, does not vary with i (i.e. $\mu_{ij} \equiv \mu_j$, as in Kingman's house-of-cards model), the effects of mutation can be described by a potential analogous to $\overline{W}$ (Wright, 1969, p. 394; Akin, 1979, pp. 64–66). In this case the non-degenerate stationary density for one locus is given by Wright's (1949) formula

$$\Psi(p) = C\overline{W}^{2N} \prod_j p_j^{(4N\mu_j - 1)}. \qquad (3.8b)$$

The extension to multiple loci in linkage equilibrium is straightforward. Wright has used this distribution extensively to give a quantitative basis to his 'shifting balance' model of evolution. Such diffusion analyses can readily be extended to give the probability of a transition from one adaptive peak to another (Wright, 1941; Gardiner, 1983, ch. 9; Lande, 1985; Barton & Rouhani, 1987). We should note, however, that this simple formulation will generally break down if recombination, frequency-dependent selection, or more complex mutation patterns are significant.

4. Genetic distance

The matrix G mediates the effects of both drift and selection. The interpretation of both processes can be simplified by regarding G as setting a natural measure of genetic distance. The distance between two populations in states that differ by a small amount δx_α is defined as $\sum\sum_{\alpha\beta} \delta x_\alpha (G^{-1})_{\alpha\beta} \delta x_\beta$. This distance is a generalization of the Mahalanobis D^2 distance. If x represents the means of a set of correlated characters, then the distance is the conventional Mahalanobis D^2, defined using the genetic covariance rather than the phenotypic covariance (Lande, 1979; Schluter, 1983). If x represents a set of allele frequencies, the distance is just Bhattacharya's (1946) and Cavalli-Sforza & Edwards' (1967) measure (see Antonelli & Strobeck, 1977). For the special case of populations whose evolution can be approximated by the mean and the variance of a quantitative character, the matrices in (5.2) and (5.3) below define a distance that takes differences in both these variables into account. In practice, differences in variance are rarely measured accurately enough for such a measure to be useful: it is proposed here more as an aid to understanding the theoretical dynamics of polygenic systems, than for the analysis of data from natural or artificial populations.

The metric g_{ij} defines a curved geometry: for example, with three alleles at a single locus, the population moves on the two-dimensional octant of a three-dimensional sphere (Antonelli & Strobeck, 1977). In the curved geometry, a set of populations that begin in the same state will diffuse out at an equal rate in all directions as a result of sampling drift; the mean square distance moved after T generations is (to leading order in $1/N$) equal to T/N (Heuch, 1975; Antonelli & Strobeck, 1977). If we choose a geodesic coordinate frame y_w in which G is diagonal, the covariance between fluctuations δy_w caused by sampling drift is simply

$$\mathrm{cov}(\delta y_w, \delta y_v) = \delta_{wv}/N, \qquad (4.1)$$

where δ_{wv} is the Kronecker delta.

The response to selection has an equally simple interpretation. When distance is measured using an arbitrary set of coordinates, a population will not necessarily move up a gradient in mean fitness as rapidly as it might; the direction it takes is the product of the additive genetic covariance and the gradient in log mean fitness (Lande, 1979; equation 1). However, if distance is measured relative to the metric G, a population will move up a gradient in log mean fitness as rapidly as possible: that is, it will move in the direction which causes the greatest increase in mean fitness over a given genetic distance. In the geodesic coordinate frame, the effect of selection is

$$dy_w/dt = \partial(\ln \overline{W})/\partial y_w. \qquad (4.2)$$

Provided fitnesses are not frequency-dependent, and the vector y_w contains all relevant variables, the above equations only require the weak selection, continuous-time approximation.

When the dynamic equations are written in tensor notation, with metric g_{ij}, they retain the same form under any choice of coordinates, i.e. they are 'covariant' (Graham, 1977). When written in this form, results on the joint effects of mutation, selection and drift can be obtained quite easily, regardless of the form of G (Graham, 1977; Barton & Rouhani, 1987; equations 4.1 and 4.2 above). However, because tensor notation is unfamiliar to most geneticists, we have not used it in this paper.

5. Application to a quantitative character

These general theoretical relations apply to weak selection on a quantitative character. This can only be exactly described by the distribution of allelic effects at all the relevant loci. We therefore seek a simple approximation by changing to a more tractable set of coordinates and finding conditions that justify neglect of all but a few of them. For simplicity, we will assume completely additive genetic effects and additive, genotype-independent environmental effects that follow a Gaussian distribution with zero mean and variance V_e. Neglect of dominance and epistasis is biologically restrictive: more general models will be considered in subsequent analyses.

Let m_i denote the average contribution of alleles at locus i, and let m_{ki} denote the kth central moment of allelic effects at this locus (i.e. $m_i = E(X_i)$ and $m_{ki} = E[(X_i - m_i)^k]$, where X_i denotes the effect of an allele at locus i and E denotes expectation with respect to the allele frequencies in the population). In Appendix B we describe the transformation from allele frequencies to these moments. The next step, which is

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

also described in Appendix B, is to relate the dynamical equations for the moments at individual loci to the observable dynamics of the overall phenotype. Our assumptions of additivity and linkage equilibrium allow us to sum over loci to obtain a description of the population in terms of the phenotypic moments, denoted $M, M_2, M_3, \ldots$. In the following equations we denote the phenotypic mean and variance by $\bar{Z} = M = 2\Sigma_i m_i$ and $V = M_2 = 2\Sigma_i m_{2i} + V_e \equiv V_g + V_e$. Suppressing the subscripts identifying individual loci, the equations for the phenotypic moments are

$$\begin{bmatrix} d\bar{Z}/dt \\ dV/dt \\ dM_3/dt \\ \vdots \end{bmatrix} \simeq 2\Sigma_i \begin{bmatrix} m_2 & m_3 & (m_4 - 3m_2^2)\ldots \\ m_3 & (m_4 - m_2^2) & (m_5 - 4m_3 m_2)\ldots \\ (m_4 - 3m_2^2) & (m_5 - 4m_3 m_2) & (m_6 - m_3^2 - 6m_2 m_4 + 9m_2^3)\ldots \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \partial(\ln \bar{W})/\partial \bar{Z} \\ \partial(\ln \bar{W})/\partial V \\ \partial(\ln \bar{W})/\partial M_3 \\ \vdots \end{bmatrix} \tag{5.1}$$

Our assumptions imply that the first two elements of the matrix in equation 5.1, namely $2\Sigma_i m_{2,i}$ and $2\Sigma_i m_{3,i}$, are simply V_g and M_3. Given our assumptions that mating is random, that linkage disequilibrium is negligible, and that selection is weak, these equations apply for any distribution of allelic effects and for any frequency-independent form of selection. Bulmer (1980, p. 174) derived the dependence of dM/dt on the $m_{3,i}$ and presented an equation (his 10.33) for dV/dt that is equivalent to the first two terms of dV/dt in eqn. 5.1. Gillespie (1984) also noted the dependence of dV/dt on $m_{3,i}$. Equation 5.1 extends these analyses to higher order moments.

Equation 5.1 shows that the evolution of the lower moments generally depends on the higher moments. In order to make further progress, two steps are necessary. First, it must be shown that the terms corresponding to selection on the phenotypic mean and variance generally dominate, so that higher-order selection coefficients (namely, $\partial(\ln \bar{W})/\partial M_3, \ldots$) can be neglected. In effect, this removes all but the first two columns of the above matrix. Second, the higher moments must be approximated by some function of the lower moments; this removes all but the first few rows of the matrix and produces a closed system of equations.

There are two alternative rationales for ignoring the higher-order selection coefficients in equation 5.1. First, under weak, Gaussian stabilizing selection (or any selection function that can be adequately approximated by the first three terms of a Taylor's series expansion), the mean fitness depends only on the mean and variance of the phenotypic distribution, irrespective of the form of this distribution. Second, for more general selection schemes, one might argue that if many loci are involved the phenotypic distribution will be approximately Gaussian, so the mean fitness will again depend only on the phenotypic mean and the variance. This argument, which is based on the central limit theorem, is delicate (see Appendix C); however, it shows that under stabilizing selection on a large number of loci, only the selection coefficients on the mean and the variance need be considered.

We next show that both the Gaussian and rare-alleles approximations impose relationships among the moments that collapse equation 5.1 into one involving only the first two or three moments at each locus. In the Gaussian model the third moment of each allelic distribution must be zero, and the fourth moment must equal three times the square of the variance. Moreover, by assumption, the phenotypic distribution must be precisely Gaussian, and so higher-order selection coefficients can be neglected even when the number of loci is small. With identical loci, equation 5.1 then simplifies to

$$\begin{bmatrix} d\bar{Z}/dt \\ dV/dt \end{bmatrix} \simeq \begin{bmatrix} V_g & 0 \\ 0 & V_g^2/n \end{bmatrix} \begin{bmatrix} \partial(\ln \bar{W})/\partial \bar{Z} \\ \partial(\ln \bar{W})/\partial V \end{bmatrix}. \tag{5.2}$$

To the extent that n is estimable, equation 5.2 achieves the goal of expressing the dynamics of polygenic variation in terms of observable parameters. However, if the loci do not have identical distributions of effects, the lower right-hand corner of the matrix becomes $4\Sigma_i m_{2i}^2$. The sum can be written as V_g^2/n_e, which defines an effective number of loci. Unfortunately, this effective number will in general change as selection changes the relative contributions of different alleles and different loci: the dynamics of the effective number depend on the genetic details that we are trying to circumvent. Hence, even the highly simplified Gaussian genetic model does not generally yield simply predictable phenotypic dynamics. The same conclusion follows from Lande's (1975) analysis of Gaussian stabilizing selection (see his equation 19).

Under the rare-alleles approximation, each locus is close to fixation. Each moment is therefore proportional to the frequency of the rare alleles, and so products of moments can be neglected in equation 5.1. Furthermore, because all the moments are proportional to the rare-allele frequencies, the fourth moment at each locus may be roughly proportional to the second moment. The following arguments support this intuition. Assuming that all loci have identical even moments, we define the ratio between the fourth and second moments to be θ. In the simplest case, where two alleles segregate at each locus, one is rare, and the effect of a substitution is 2α at each locus, $\theta = \alpha^2$. For the continuum-of-alleles model analysed by Turelli

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

(1984), in which mutations have a Gaussian distribution of phenotypic effects with variance α^2, and stabilizing selection follows a Gaussian fitness function, the house-of-cards (low mutation rate) approximation also leads to $\theta = \alpha^2$ (see equations 3.14, 3.15 in Turelli, 1984). Although the fourth and second moments will not be strictly proportional under general selection schemes with multiple alleles, the applicability of this approximation to the two special cases considered above suggests that it may be reasonable whenever rare alleles contribute most of the genetic variance. In general, we expect that when most genetic variation is attributable to rare alleles with effects large relative to $\sqrt{(m_{2,t})}$, $m_4 - m_2^2 \simeq m_4 \simeq \theta m_2$, where θ will be a small factor of the same order as the square of the effect of the alleles; its value will depend on the genetic details. Consideration of diallelic loci motivates two additional simplifying assumptions: that $m_5 \simeq \theta m_3$, and $m_6 \simeq \theta^2 m_2$ (under the continuum-of-alleles model, $m_5 = m_3 = 0$, and $m_6 \simeq 3\theta^2 m_2$). These approximations lead to

$$\begin{bmatrix} d\bar{Z}/dt \\ dV/dt \\ dM_3/dt \end{bmatrix} \simeq \begin{bmatrix} V_g & M_3 & \theta V_g \\ M_3 & \theta V_g & \theta M_3 \\ \theta V_g & \theta M_3 & \theta^2 V_g \end{bmatrix} \begin{bmatrix} \partial(\ln \bar{W})/\partial \bar{Z} \\ \partial(\ln \bar{W})/\partial V \\ \partial(\ln \bar{W})/\partial M_3 \end{bmatrix}. \quad (5.3)$$

(Here we have included the third-order selection coefficient for completeness; however, the arguments in Appendix C suggest that in most cases the third and higher columns of the matrix will be negligible for large n.) Note that $dM_3/dt \simeq \theta d\bar{Z}/dt$. If M_3 remains near zero, equation 5.3 can be reduced to a two-moment approximation that differs from equation 5.2 only in replacing V_g^2/n by θV_g in the second diagonal term. Whereas equation 5.2 requires an estimate of the number of loci, 5.3 requires an estimate of $\alpha = \sqrt{\theta}$, the average phenotypic effect of an individual mutation.

For non-identical loci, the dynamics of the phenotypic moments can be expressed in closed form only if additional assumptions are made. For instance, if we can assume that α^2 is uncorrelated with m_2 and m_3 across loci, equation 5.3 still applies, with θ replaced by the average of α^2 across loci. At a symmetric mutation–selection equilibrium, this is equivalent to assuming that mutation rates are uncorrelated with average effects of mutations. However, away from equilibrium, changes in the variance produced by response to selection will certainly be correlated with allelic effects (α). Thus it is unclear whether approximation 5.3 is any more general than 5.2, which requires identical loci. Nevertheless, we will see from the numerical examples below that equation 5.3 can successfully accommodate some differences among loci.

An important feature of equation 5.3 is the role played by the skew term, M_3 (see Bulmer, 1980, and Gillespie, 1984, for a similar result). In previous analyses of mutation and stabilizing selection (e.g. Wright, 1935; Kimura, 1965; Bulmer, 1972; Latter, 1960), the models have been symmetric, and the mean has been assumed to lie at the optimum. This ensures that M_3 is zero at equilibrium. Although the observed near-normality of most quantitative characters (Falconer, 1981) suggests that the third moment is not very large, it is not clear that it is negligible. When mutation is added to the model (see Section 6 below), equation 5.3 is modified by the incorporation of mutational variance, but the dynamics of the mean and skew are unaltered. At equilibrium, the system of three simultaneous equations becomes degenerate (because $dM_3/dt \simeq \theta d\bar{Z}/dt$), and the equilibrium value of M_3 is a free parameter. In Wright's (1935) model, in which two alleles segregate at each of n interchangeable loci, Barton (1986) found that many stable equilibria are possible under mutation and stabilizing selection. These differ in the magnitude of M_3, as expected from equation 5.3. The third moment is negligible only if the population mean lies near the optimum phenotype. In general, equation 5.3 will only give closed equations for the evolution of the mean and the variance if some process, such as sampling drift, ensures that the third moment remains small. This complex question is discussed in more detail in Barton (1986).

6. Mutation and stabilizing selection: equilibria

To show that our general formulation is consistent with previous work and allows it to be extended, equations 5.2 and 5.3 can be applied to find the variance maintained at an equilibrium between mutation and stabilizing selection. Following the analyses of Lande (1975) and Turelli (1984), we assume Gaussian stabilizing selection. The fitness of an individual with phenotype Z is

$$W(Z) = \exp[-(Z - Z_0)^2/2V_s]. \quad (6.1)$$

Here Z_0 denotes the optimal phenotype, and $1/V_s$ measures the intensity of selection. Assuming a Gaussian distribution of phenotypes, $p(Z)$, with mean $\bar{Z}$ and variance V, calculation of $\bar{W} = \int p(Z)\, W(Z)\, dZ$ gives

$$\ln \bar{W} = -\tfrac{1}{2}[\ln(1 + V/V_s) + (\bar{Z} - Z_0)^2/(V + V_s)]. \quad (6.2)$$

If selection is relatively weak and the population mean is not too far from the optimum (i.e. $V_s \gg V$, $(\bar{Z} - Z_0)^2$), we have

$$\partial(\ln \bar{W})/\partial \bar{Z} \simeq -(\bar{Z} - Z_0)/V_s \quad \text{and}$$
$$\partial(\ln \bar{W})/\partial V \simeq -\tfrac{1}{2}V_s. \quad (6.3)$$

If μ denotes the mutation rate, and α^2 the variance of allelic effects produced by mutation at each locus, mutation can be incorporated into our general approximation (5.1) by adding $V_m = 2n\mu\alpha^2$ to the equation for dV/dt. It is, of course, possible that mutation alters the mean and the skew; however,

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

because we have no reason to expect mutation to act in any particular direction, we ignore this possibility (see Appendix C).

Under the Gaussian genetic model with weak selection, equations 5.2 and 6.3 predict that the equilibrium phenotypic mean is Z_0 and the equilibrium genetic variance is

$$V_g \simeq \sqrt{(2nV_m V_s)} = \sqrt{(4n^2\mu\alpha^2 V_s)}. \quad (6.4)$$

This agrees with the Gaussian-based results of Kimura (1965), Latter (1970) and Lande (1975) if the small contribution of linkage disequilibrium to Lande's prediction is ignored (see equation 3.3 of Turelli, 1984). By analysing each locus separately (as in equation B 3), we can recover the more general results given by Kimura and Lande for non-identical loci.

Under the rare-alleles approximation with weak selection, equations 5.3 and 6.3 imply that the equilibrium $\overline{Z}$ is the optimum Z_0 only if M_3 is zero. When it is, we recover the standard result for the equilibrium genetic variance,

$$V_g \simeq 2V_m V_s/\alpha^2 = 4n\mu V_s \quad (6.5)$$

(cf. Latter, 1960; Bulmer, 1973; Turelli, 1984). Thus in our analysis the differences between the Gaussian and house-of-cards (rare alleles) predictions emerge from the different constraints that they impose on the higher moments in equation 5.1. Prediction 6.5 agrees with previous results for low per-locus mutation rates and generalizes them to an arbitrary number of alleles per locus, under the assumption that most phenotypic variance is due to rare alleles of relatively large effect. By considering each locus separately, this prediction can be extended to non-interchangeable loci.

As noted above, M_3 is a free parameter under the rare-alleles approximation (5.3). When the third moment is not negligible at equilibrium, the variance is higher than predicted by equation 6.5. Using Wright's (1935) model of multiple diallelic loci, Barton (1986) discovered stable mutation–selection equilibria at which the mean did not lie at the optimum and the equilibrium skew was non-zero. By using the skew observed at numerically determined asymmetric equilibria, he found that the observed variance agreed well with the prediction from equation 5.3.

7. Mutation and stabilizing selection: dynamics

To examine the dependence of phenotypic evolution on the underlying genetic structure, we will consider a population initially at equilibrium between mutation and Gaussian stabilizing selection with $M_3 = 0$. Suppose that the optimum phenotype, Z_0, now changes to a new value. Equations 5.1 and 6.3 imply that for weak selection (i.e. $V_s \gg V$, $(\overline{Z}-Z_0)^2$), the character will initially change according to

$$d\overline{Z}/dt \simeq -V_g(\overline{Z}-Z_0)/V_s, \quad (7.1a)$$

$$dV/dt \simeq 0, \quad (7.1b)$$

irrespective of the form of the distribution of allelic effects, and whether or not the loci are identical. Within each generation, selection changes the mean by $-(\overline{Z}-Z_0)V/(V+V_s) \simeq -(\overline{Z}-Z_0)V/V_s$, so the prediction for the mean (7.1*a*) agrees with the classical quantitative genetic equation for selection response and Lande's (1976) Gaussian phenotypic analysis. The additional prediction that the variance remains approximately constant supports a frequent assumption of these phenotypic analyses. If evolution proceeds without producing significant skew in the distribution of the character, equations 7.1 continue to hold and the underlying genetic details are irrelevant.

The question now is, will enough skew be generated by the shift in optimum to change the genetic variance? Under the Gaussian genetic approximation, skew is absent by assumption and none can be generated (5.2). On the alternative rare-alleles approximation, (5.3) predicts that skew will be generated. Denote the deviation from the new optimum by $\Delta = \overline{Z}-Z_0$ and denote the initial equilibrium variance with $M_3 = 0$ by V_0. We will assume that $m_4 - m_2^2 \simeq \alpha^2 m_2$, and that α^2 is of order $1/n$. Then including only terms of order $1/n$ (so that we can ignore $\partial(\ln \overline{W})/\partial M_3$, etc.), equations 5.3 and 6.3 produce

$$d\Delta/dt \simeq -\Delta V_g/V_s - M_3/2V_s \quad (7.2a)$$

$$dV/dt \simeq -\Delta M_3/V_s - \alpha^2 V_g/2V_s + V_m \quad (7.2b)$$

$$dM_3/dt \simeq -\alpha^2\Delta V_g/V_s - \alpha^2 M_3/2V_s \simeq \alpha^2 d\Delta/dt. \quad (7.2c)$$

Equation 7.2*c* implies that if the skew is initially zero, $M_3 \simeq \alpha^2(\Delta-\Delta_0)$. Substituting this into equations 7.2*a* and 7.2*b* yields

$$d\Delta/dt \simeq -\Delta V_g/V_s - \alpha^2(\Delta-\Delta_0)/2V_s \quad (7.3a)$$

$$dV/dt \simeq \alpha^2\Delta(\Delta_0-\Delta)/V_s - \alpha^2(V-V_0)/2V_s. \quad (7.3b)$$

In deriving equation 7.3*b*, we have used our assumption that $M_3 = 0$ initially so that the initial genetic variance, $V_{g,0}$, satisfies $V_m = \alpha^2 V_{g,0}/2V_s$, and our assumption that V_e is constant so that the V and V_g have identical dynamics.

Equations 7.3 show that the response to the change in optimum occurs over two timescales. First, the mean rapidly moves towards the new optimum, generating skew and variance. To obtain relatively simple analytical predictions from equation 7.3, we introduce the additional assumption that $\alpha^2 \ll V_{g,0} \equiv 2nm_2$. The number of loci, n, must be large, say on the order of 100, for this assumption to be consistent with the house-of-cards assumption $\alpha^2 \gg m_2$. If $\alpha^2 \ll V_{g,0}$, the first term on the right-hand side of 7.3*a* will dominate while Δ is still much larger than $\alpha^2\Delta_0/V_g$, and V will change relatively little. Thus we have the approximate solution

$$\Delta \simeq \Delta_0 \exp(-V_{g,0}\, t/V_s). \quad (7.4a)$$

Because the mean is changing faster than the variance, the first term on the right-hand side of equation 7.3*b*

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

dominates during this period. Substituting approximation 7.4*a* into 7.3*b* produces

$$V \simeq V_0+(\Delta_0^2\alpha^2/2V_{g,0})[1-\exp(-V_{g,0}t/V_s)]^2. \quad (7.4b)$$

Thus the variance will be increased by approximately $\Delta_0^2\alpha^2/2V_{g,0}$ during this period of relatively rapid change.

Once Δ is small, subsequent changes are much slower. If we now take the leading terms in equation 7.3, on the assumption that Δ is of order α^2, we obtain

$$\Delta \simeq \alpha^2\Delta_0/2V_{g,0}+0(\alpha^4) \quad (7.5a)$$

$$dV/dt \simeq \alpha^2\Delta\Delta_0/V_s-\alpha^2(V-V_0)/2V_s$$

$$\simeq \alpha^4\Delta_0^2/(2V_{g,0}V_s)-\alpha^2(V-V_0)/2V_s. \quad (7.5b)$$

Hence,

$$V \simeq V_0+(\alpha^2\Delta_0^2/V_{g,0})[1-\tfrac{1}{2}\exp(-\alpha^2t/2V_s)]. \quad (7.6)$$

This shows that a second increment of $\Delta_0^2\alpha^2/2V_{g,0}$ is slowly added to V after the initial rapid increase described by equation 7.4*b*. Overall, the proportional increase in the genetic variance is $(\Delta_0\alpha/V_{g,0})^2$. If we relax the requirement that $\alpha^2 \ll V_{g,0}$, equations 7.3 are more difficult to analyse, but they always predict that V will increase.

The exact dynamics of the response of a polygenic system to a change in optimum is complex, but consistent with these approximations for appropriate parameter values (see Barton, 1986, and the numerical results in the following section). Approximations 7.4 and 7.5 assume that the alleles responsible for variation remain rare and that $\alpha^2 \ll V_{g,0}$. They imply that the effects of the skew on the variance will be small only if

$$\Delta_0/\sqrt{(V_{g,0})} \ll 1/(\alpha/\sqrt{V_{g,0}}), \quad (7.7)$$

i.e. the shift in optimum must be smaller than the inverse of the effect of mutations at individual loci (both being expressed relative to the standard deviation of the initial distribution of genetic effects). This places constraints on both the number of loci and the magnitude of phenotypic change that are consistent with small increases in genetic variance under the rare-alleles approximation.

It seems reasonable to conjecture that the variance dynamics seen in the Gaussian and rare-alleles approximations bracket the behaviour of less extreme models. The increase in variance is driven by skew developing in response to selection on the mean. From equation 5.1 it follows that the leading term in the change of M_3 is proportional to $2\Sigma_i(m_{4,i}-3m_{2,i}^2)$. Under the Gaussian model there is no kurtosis, i.e. $m_{4,i}-3m_{2,i}^2 \equiv 0$ at each locus, and so there is no tendency for skew to develop. This leads to the prediction that if the variance is initially at equilibrium, it will not change in response to directional selection. In contrast, the rare-allele approximation assumes extreme kurtosis, so that $m_4-3m_2^2 \simeq m_4 \simeq \alpha^2 m_2$ at each locus. We suspect that this leads to the maximum possible increase of variance in response to selection. This conjecture is consistent with Slatkin's (1987) results for an approximation that bridges these two extremes.

8. Numerical examples

Various calculations were performed to test the accuracy of the rare-alleles predictions concerning the response to a shift in the optimal phenotype. We have not checked the Gaussian predictions because they follow directly from properties of the Gaussian distribution if mutation rates are high enough relative to selection intensity to justify this approximation (see Turelli, 1984). A systematic survey of alternative models and parameter values would be a major enterprise; here we present a sample of numerical results that illustrates both the strengths and weaknesses of our approximations.

(i) *Numerical methods and choice of parameter values*

Because we are concerned with characters affected by many loci, deterministic calculations involving gamete frequencies are not feasible. Thus we have concentrated on calculations that assume global linkage equilibrium, so that only allele frequencies need be monitored. Another approximation, which simplifies our calculations, involves the computation of one-locus fitnesses. In general, the induced fitness of a genotype at locus 1, $W^*(G_1)$, is the average fitness of individuals with the specified genotype at locus 1, i.e.

$$W^*(G_1) = E[W(G_1, G_2, \ldots, G_n)|\, G_1], \quad (8.1)$$

where G_i denotes a genotype at locus i and the conditional expectation is taken over the genotype frequencies in the population. Assuming global linkage equilibrium and Gaussian stabilizing selection (i.e. equation 6.1), a Taylor's series expansion about $(E(G_1), E(G_2), \ldots, E(G_n))$ produces

$$W^*(G_1) = \exp\{-[G_1-E(G_1)+(\bar{Z}-Z_0)]^2/2V_s\} + O(V/V_s). \quad (8.2)$$

Thus for weak selection, i.e. $V_s \gg V$, $(\bar{Z}-Z_0)^2$, the induced one-locus fitnesses can be approximated by the leading term in (8.2). This approximation produces the same allele frequency recursions as the approximations presented in the previous two sections.

Using these approximations, we iterated allele frequencies according to the standard one-locus selection recursion (equation 3.3). Calculations were performed with two, three, four or five alleles at each of 100 loci. At a locus with s alleles, the alleles were assigned additive effects $0, \alpha_i, 2a_i, \ldots, (s-1)\alpha_i$. The parameter α_i was allowed to vary slightly across loci to avoid artifacts associated with excessive symmetry. For diallelic loci we assumed equal mutation rates between the alleles. For multiallelic loci the extreme alleles mutated only to their 'nearest neighbour' at

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

rate μ_i, whereas 'central' alleles mutated either up or down one step at rate $\mu_i/2$. The mutation rate was also allowed to vary slightly across loci. Similar models have been analysed by Chakraborty & Nei (1982) and Slatkin (1987). For each set of calculations, eight parameters were specified: s, the number of alleles per locus; V_s, the stabilizing selection parameter; $E(\alpha^2)$, the average variance of effects associated with new mutations; $CV(\alpha^2)$, the coefficient of variation of α^2 across loci; $E(\mu)$ and $CV(\mu)$, the mean and coefficient of variation of per-locus mutation rates; the initial optimum; and Δ_0, the change in the optimum. The α_i^2 and μ_i were drawn from independent lognormal distributions. For all of the calculations detailed here, $V_s = 20$, $E(\alpha^2) = 0{\cdot}05$, $E(\mu) = 10^{-4}$ and $CV(\alpha^2) = CV(\mu) = 0.1$. These parameters are scaled relative to $V_e = 1$. (See Turelli (1984) for a discussion of these values.) Other parameter values were also used, and some results are briefly described. The initial phenotypic optimum was always $(s-1)\Sigma_i \alpha_i$, which is the midpoint of the range of possible phenotypes.

For each set of parameters, we first iterated the allele frequency recursions to find a stable mutation–selection equilibrium. The initial conditions for these iterations were randomly selected allele frequencies near those expected at symmetric (i.e. $M_3 = 0$) equilibria. Allele frequencies were iterated with mutation following selection, and the moments were calculated before selection. Once a new optimum was chosen, we monitored the dynamics and equilibria of allele frequencies and the first four genotypic moments. The observed dynamics were compared to the analytical predictions based on equation 7.4 and to numerical predictions based on iterating the discrete-time analogue of equation 7.2 (called the 3-moment predictions) and the corresponding recursions with $M_3 \equiv 0$ (called the 2-moment predictions). The initial conditions for these recursions were the numerically determined initial equilibrium values. Observed equilibria after directional selection were compared to the equilibria produced by iterating the moment equations. Our empirical criterion for equilibrium was that the moments did not change to five significant digits for a period equal to that needed to reach the final values. Most of the runs were continued for 100000 generations. All the calculations were performed in double precision on a Symmetric 375 microcomputer (providing approximately 16 accurate decimal digits).

To check the accuracy of the approximations used in the 100-locus calculations, we examined the transient behaviour of six diallelic loci with a simple modification of the program that produced the six-locus equilibrium results reported in Turelli (1984). The simulations monitored gamete frequencies for six unlinked loci with $\mu_i = 10^{-4}$ and $\alpha_i^2 = 0{\cdot}05$ for each locus and $V_s = 20$. Iterations were begun with allele frequencies near a symmetric mutation–selection equilibrium corresponding to $\Delta_0 = 0$. At equilibrium, $\sqrt{V_g} \simeq 0{\cdot}22$, as expected from the rare-alleles approximation. The optima were then displaced by $\Delta_0 = 0{\cdot}02$, $0{\cdot}45$ and $0{\cdot}7$. We compared the actual mean and variance for the additive genetic effects with those predicted from iterating the one-locus allele frequency recursions based on linkage equilibrium and the fitness approximation (8.2). The percentage relative errors [i.e. $100 \times$ (predicted − observed)/observed] of the predictions for the mean were always less than $0{\cdot}5\%$. With $\Delta_0 = 0{\cdot}2$, the variance increased monotonically from $4{\cdot}8 \times 10^{-2}$ to $6{\cdot}4 \times 10^{-2}$, and throughout the predictions erred by less than $0{\cdot}5\%$. For $\Delta_0 = 0{\cdot}45$ and $0{\cdot}7$, V_g increased to above $9{\cdot}0 \times 10^{-2}$ while the mean shifted, then dropped to a value near the initial equilibrium. The variance predictions remained within $0{\cdot}5\%$ of the actual values during the increase and at equilibrium, but erred by as much as -7% during the decline. We interpret these results as supporting our simplifying assumptions.

(ii) *Dynamics of the mean*

For 100 loci with $V_s = 20$, $E(\mu) = 10^{-4}$ and $E(\alpha^2) = 0{\cdot}5$ (thus, α_i close to $0{\cdot}22$), the house-of-cards approximation predicts an equilibrium genetic variance of approximately $0{\cdot}8$, assuming $M_3 = 0$. We examined equilibria and dynamics using two, three, four and five alleles per locus and Δ_0 equal to 1, 2, 3 and 4. Note that $\Delta_0 = 1$ corresponds to approximately $1{\cdot}1$ genetic standard deviations; and if $V_e = 1$, this corresponds to about $0{\cdot}75$ phenotypic standard deviations. The results concerning the dynamics of the mean can be easily summarized. Over all the cases examined, the percentage relative errors of both the 2-moment and 3-moment recursions were always less than $0{\cdot}5\%$. The reason the 2-moment recursions remain accurate, even though the variance increased by more than 50% in some cases, is revealed by approximation 7.4*a*: most of the change in the mean occurs before there is a significant increase in the variance. The residual change in the mean is small. Figs. 2*a* and 2*b* display two examples, both with $\Delta_0 = 2$. To illustrate the two time scales of change discussed in the previous section, the time axis is logarithmic (with the initial generation set to 1). The first example uses diallelic loci, the second loci with five alleles. These were chosen because in the first case the rare-alleles approximation does not accurately predict the dynamics of the variance, whereas in the second the 3-moment predictions fit reasonably well. As the figures show, both the 2-moment and 3-moment recursions predict the dynamics of the mean phenotype very accurately.

(iii) *Dynamics of the variance*

Figs. 2*c* and 2*d* present the transient behaviour of the genetic variance for the two cases discussed above. They are qualitatively different. The two-allele model

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

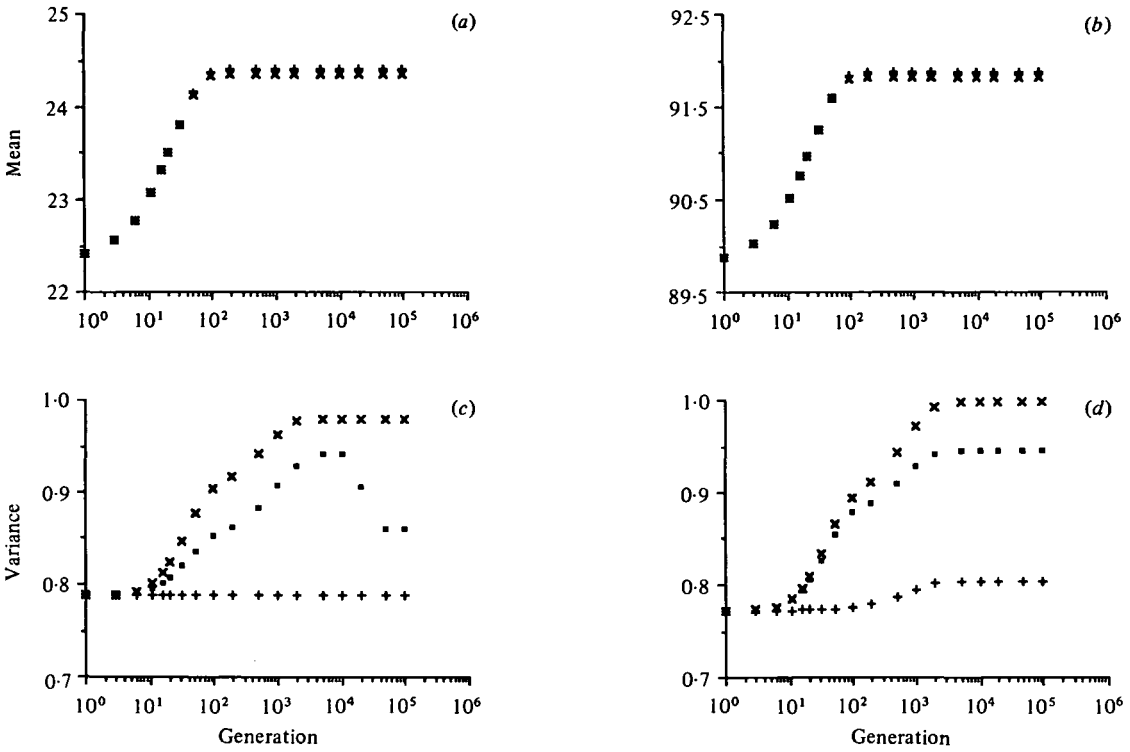

Fig. 2. The dynamics of the mean and genetic variance of a quantitative character subject to a shift in phenotypic optimum of approximately 2·2 genetic standard deviations ($\Delta_0 = 2$). Variance in the character is determined by 100 loci, with either two or five alleles segregating at each locus. ■, Numerically determined values from the 100-locus recursions; ×, predictions from the 3-moment rare-alleles recursions; +, the 2-moment predictions. The parameter values and calculation methods are described in the text.

shows a temporary increase in the variance that agrees roughly with the predictions of the 3-moment recursion. However, after the initial increase the variance falls back towards the value at the initial equilibrium. The initial increase, which is predicted by our analysis, is caused by a rapid change in allele frequencies that shifts the mean towards the new optimum. In this example the allele frequencies become large enough to approach an unstable equilibrium. They then shift slowly to a new stable equilibrium, involving substitutions at several of the loci, without significantly changing the mean. These substitutions are not accounted for in our analysis, which assumes that the alleles producing genetic variance always remain rare. This highlights a crucial assumption. Our approximation that variance-producing alleles are rare implies that changes in the mean phenotype are accomplished by small changes in the frequencies of many alleles rather than by substitutions at a few loci. Clearly, this requires large numbers of loci and only moderate shifts in the optimum. For instance, with $\Delta_0 = 1$, the diallelic model behaves as predicted: the variance remains high after the initial increase (see Table 2 below).

Fig. 2*d* shows the dynamics of the variance in a population identical to that represented in Fig. 2*c* except that five alleles segregate at each locus. In this case the variance remains high at the new equilibrium. This shows that the conditions under which the rare-alleles approximation collapses are sensitive to genetic details such as the number of segregating alleles. Even in this example the equilibrium is not quite as high as predicted by the 3-moment recursions because alleles do not remain sufficiently rare to justify ignoring products of allele frequencies. Another apparent anomaly is the observable increase in the 2-moment variance prediction. This seems to contradict the conclusion of the previous section that the variance prediction should remain constant, at its equilibrium value, if skew is not taken into account. The reason for the discrepancy is that the initial conditions for the predictions correspond to the *observed* initial equilibrium values rather than the initial equilibrium values predicted by the 2- and 3-moment recursions.

Fig. 2 illustrates a dichotomy between the relative robustness of our variance predictions while the population mean is changing and the difficulty of predicting equilibrium variances after directional selection has ended. Equations 7.4 and 7.5 predict that once the population mean has shifted to within roughly $\Delta_0 \alpha^2 / V_{g,0}$ of the new optimum, the proportional increase of the genetic variance will be approximately $\frac{1}{2}(\Delta_0 \alpha / V_{g,0})^2$. For the parameters we used, the population mean settles near the new optimum in about 100 generations (Fig. 2). Table 1

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

Table 1. *Proportional increases in the genetic variance after 100 generations of response to selection and, in parentheses, the analytical predictions based on iterating equation 7.2 for different numbers of alleles per locus and different initial displacements of the phenotypic optimum* (Δ_0)

	Δ_0			
Allele	1	2	3	4
2	0·02 (0·04)	0·08 (0·16)	0·17 (0·36)	0·27 (0·64)
3	0·03 (0·05)	0·13 (0·18)	0·27 (0·41)	0·45 (0·72)
4	0·03 (0·04)	0·13 (0·17)	0·27 (0·39)	0·45 (0·69)
5	0·04 (0·04)	0·14 (0·17)	0·29 (0·38)	0·47 (0·67)

All of these runs assume 100 loci with global linkage equilibrium. The per-locus mutation rates, μ_i, and squared allelic effects (spacings), α_i^2, are independent, lognormally distributed random variables with $E(\mu) = 10^{-4}$, $E(\alpha^2) = 0{\cdot}5$ and $CV(\mu) = CV(\alpha^2) = 0{\cdot}1$

Table 2. *Proportional increases in the genetic variance after equilibration to a new phenotypic optimum and, in parentheses, the percentage relative errors of the 3-moment variance predictions based on iterating equation 7.2*

	Δ_0				
Allele	0*	1	2	3	4
2	(0)	0·06 (1)	0·09 (14)	0·03 (43)	0·01 (70)
3	(8)	0·07 (9)	0·23 (10)	0·34 (18)	0·13 (65)
4	(6)	0·06 (7)	0·22 (8)	0·33 (16)	0·28 (41)
5	(4)	0·07 (4)	0·22 (6)	0·26 (21)	0·21 (47)

* The parenthetic values in this column are the percentage relative errors of the symmetric house-of-cards approximation for the initial equilibrium variance, i.e. $V_g \simeq 4\Sigma_i \mu_i V_s$.

summarizes the observed proportional increases in the variance after 100 generations and the analytical predictions based on iterating equation 7.2. Although the predicted increases are systematically higher than the values observed, the observed increases are proportional to Δ_0^2, as predicted. Moreover, the increases seen with three, four and five alleles are fairly consistent.

(iv) *Equilibria reached after a population shift*

Table 2 presents the proportional increases in the genetic variance after the new equilibrium has been reached, and the percentage relative errors of the 3-moment variance predictions based on iterating equation 7.2. For comparison, the first column gives the percentage relative error of the house-of-cards prediction for the initial equilibrium. For $\Delta_0 = 1$, final equilibrium predictions are as accurate as the initial equilibrium predictions for all four genetic models. For $\Delta_0 = 2$, only the 2-allele model shows an ultimate increase in the variance that is not approximately twice as large as that after 100 generations (cf. Table 1 and Fig. 2*c*). However, for $\Delta_0 = 3$ and 4, the behaviour of all four models departs substantially from our predictions: the magnitude of the departure depends on the model.

For these examples, $1/\sqrt{(V_{g,0})} \simeq 1{\cdot}1$ and $1/(\alpha/\sqrt{V_{g,0}}) \simeq 4$. Thus our condition (7.7) for the change in the variance to be small, i.e. $\Delta_0/\sqrt{V_{g,0}} \ll 1/(\alpha/\sqrt{V_{g,0}})$, requires that Δ_0 be much less than 4 (say <2). Clearly our approximations are quite accurate for such small changes. The results above suggest that for the parameters used, the approximations remain accurate even when predicting proportional increases of the genetic variance up to 10–20%. Larger increases can be accurately predicted if the sizes of the allelic effects are increased. For instance, with $E(\alpha^2) = 0{\cdot}8$, numerical iterations of the 5-allele model show that the genetic variance more than doubles (from 0·81 to 1·91) over 100 generations in response to an optimum shift of $\Delta_0 = 2$. Because

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

these alleles of large effect remain rare, the maximum relative error of the 3-moment predictions is only 2% throughout the course of selection response. Although this agreement between theory and numerical observation supports our approximations, there is little empirical support for such dramatic increases of genetic variance in response to directional selection. This illustrates the constraints on parameter values for which the rare-alleles analysis is not only mathematically accurate but also makes biologically realistic predictions.

9. Discussion

Provided selection is weak relative to recombination, the effects of frequency-independent selection and drift can be described by a positive definite matrix and the gradient of log mean fitness. Such a description can be expressed in a form that is independent of the particular variables used to describe the population. This description rests on the idea that the covariance matrix establishes a natural measure of genetic distance, and hence a natural geometry for selection and drift. For the particular case of an additive polygenic character, it provides a convenient way of transforming the coordinates from the distribution of allelic effects at each locus, to the moments of these distributions. In general, this new formulation is no simpler; however, under two alternative approximations, the Gaussian and the rare-alleles, the description of the dynamics collapses into one involving only the first two or three moments at each locus. If the loci have identical distributions of effects, and if the third moment of the phenotypic distribution is negligible, then both of these approximations lead to a pair of closed equations for the mean and the variance. These equations apply under any arbitrary selection scheme and, in the particular case of a balance between mutation and stabilizing selection, predict the same variance as more specialized models.

These simple approximations to the dynamics of the mean and variance may hold under only a very restricted range of conditions. The aim of this paper has been not so much to derive an approximation that will successfully explain experimental data, but rather to set out a general framework that delimits the applicability of various possible approximations. We feel that any comparison between the approximations set out here and the real world can only be tentative. At the least, selection on more than a single character and pleiotropy must be considered. Bearing this cautionary note in mind, we will make some comments on the applicability of our results.

The critical issue is whether there is significant change in the genetic variance during response to directional selection. Under the Gaussian genetic model, the distribution of allelic effects never becomes skewed (by assumption); hence selection on the mean does not affect the genetic variance. However, if the genetic variance is not at equilibrium, or if selection acts directly on the variance (e.g. through changes in the intensity of stabilizing selection), the dynamics of the variance are predictable from phenotypic observations only if all loci have identical distributions of effects and the number of loci is known. If this condition is not satisfied, one could still define an 'effective number of loci' (see Section 5 above); but this would itself evolve under selection, and so the model would lose much of its predictive power.

Under the rare-alleles approximation, evolutionary dynamics can be summarized in terms of observable variables only if an estimate of the average effects of individual mutations is available, and if allelic effects are uncorrelated with genetic variances and skews at individual loci. Whether this second condition can be satisfied by loci with very different effects remains to be determined.

Equations 5.2 and 5.3 can easily be generalized to describe the evolution of the phenotypic means and covariances for a set of correlated characters. However, the sufficient condition for a phenotypic description, that the distribution of allelic effects are equally at all loci, becomes much more restrictive; the distributions of the *vectors* of effects on all the correlated characters must be identical for all loci (Turelli, 1985). Approximations for selection on correlated characters are complex. However, because of the dramatic effects of pleiotropy and correlation on mutation–selection equilibrium predictions (Turelli, 1985) and the dynamics of selection (Lande & Arnold, 1983), we feel that the analysis of multicharacter selection is necessary before one attempts to relate the predictions of alternative models to data. For now, the rare-alleles approximation might be viewed as an extreme alternative against which to test the robustness of predictions derived from Gaussian genetic analyses.

The recursions for identical loci reveal some important differences between the dynamics of the genetic variance expected under the Gaussian and rare-alleles approximations. According to the Gaussian approximation, the variance responds to selection on the variance (i.e. $\partial(\ln \overline{W})/\partial V$) in proportion to V_g^2/n (see equation 5.2). In contrast, the rare-alleles approximation (see equation 5.3), which is based on the assumption that α^2 is much larger than V_g/n, predicts that the variance will respond in proportion to $\alpha^2 V_g$. Thus if per-locus mutation rates are sufficiently low, and selection sufficiently strong, to justify the rare-alleles approximation, the genetic variance is expected to change much more rapidly. A more fundamental difference concerns the connexion between the dynamics of the mean and variance. As shown by the general equations 5.1, selection on the mean can influence the genetic variance only if the distribution of allelic effects is skewed. The Gaussian model rules this out by assumption. Among other things this leads to the conclusion that under

stabilizing selection, fluctuations in the optimal phenotype will have no influence on the equilibrium genetic variance (Lande, 1977). The coupling of the dynamics of the mean and variance under models that allow skew raises the possibility that fluctuating optima can increase the equilibrium level of genetic variation. We are currently investigating this possibility. In general there seems to be little reason to concentrate exclusively on mutation–selection equilibrium as a mechanism maintaining polygenic variation.

The range of applicability of the approximations suggested here can only be determined by simulations of a range of genetic models, and a range of selection schemes. However a general argument, based on the method developed above, suggests that no general approximation is likely to be able to describe a polygenic system far from equilibrium solely in terms of measurable, phenotypic variables. Imagine a distortion of the coordinate system, on which the measurable parameters change along the vertical axis, and the remaining coordinates change along the horizontal axis. Under selection and drift, an ensemble of populations will occupy some distribution across this space, which will evolve through time. We might hope to average over the unmeasurable horizontal coordinates, and so derive equations for the evolution of the coordinates of interest. (This procedure, which arises in an essentially identical context in statistical thermodynamics, is known as 'adiabatic elimination', see Gardiner, 1983.) However, this will only be possible if the distribution across the transverse coordinates takes on a characteristic form that is independent of the evolutionary history of the population. Such independence may be approximated in special cases, such as those considered here, but seems unlikely in general. Even in Wright's (1935) simple polygenic model with diallelic loci, multiple equilibria are possible, so that the state of the population depends strongly on its previous history (Barton, 1986). More generally, selection pressures on the transverse variables are likely to be weaker than those on the character of interest, and so the transverse distribution is likely to change more slowly than this character, and will depend on the history of the population as well as on the current mean and variance. It may be that the evolution of the genetic variance depends so much on the detailed genetic basis of the character that it is essentially unpredictable.

This work was supported by grants from the NSF and the Guggenheim Foundation (to M. Turelli), and the SERC (to N. H. Barton). We would like to thank J. A. Coyne, J. Felsenstein, M. Kirkpatrick, L. Mueller, S. Rouhani, M. Slatkin, J. B. Walsh and two anonymous reviewers for their helpful comments.

Appendix A

Interpretation of partial derivatives

The derivative in equation 3.4 can be interpreted in two ways. We can explicitly impose the constraint that the allele frequencies sum to one, so that only $s-1$ variables are monitored (Fig. 1*b*), or we can ignore this constraint and treat all allele frequencies independently (Fig. 1*a*). In the latter case the mean fitness, $\overline{W}$, must be regarded as a function that could, in principle, range over all values of p_i, including those where the allele frequencies do not sum to one. Although this is biologically nonsensical, it is mathematically convenient. The differential is then taken along the direction of increasing p_i; this will not lie in the plane where $\Sigma_i p_i = 1$. However, one can easily verify that the form of G ensures that the sum of p_i never changes under selection, so that the population will always remain in the real world: in other words, there is no genetic variance out of the plane $\Sigma_i p_i = 1$. Wright used yet another method: he wrote equation 3.4 as $\Delta p_i = \frac{1}{2} p_i(1-p_i)\,\partial(\ln \overline{W})/\partial p_i$, but used a partial differential defined along the direction in which p_i changes, but the relative proportions of all the other alleles remain constant (Wright, 1969, pp. 38–39).

Appendix B

(i) *Transformation from discrete allele frequencies to moments*

To predict the dynamics of phenotypic moments, we begin by finding the dynamics of the moments of the allelic effects at an individual locus. This is calculated by transforming the matrix G (equation 3.1) from a representation in terms of allele frequencies to one in terms of moments of allelic effects. We begin by working through the familiar case of discrete alleles, and then give a more general derivation, based directly on the (possibly continuous) distribution of allelic effects. Let X_i denote the effect of an allele at locus i, let $m_i = E(X_i)$ denote its mean, and $m_{ki} = E[(X_i - m_i)^k]$ for $k = 1, 2, \ldots$. To simplify the notation, we will temporarily concentrate on a specific locus, and suppress the subscript that identifies it. Let p_j denote the frequency of allele j at this locus, and let x_j denote its phenotypic effect. Using the convention that for a locus with s alleles, $p_s = 1 - \sum_{j=1}^{s-1} p_j$, it follows that for $j = 1, \ldots, s-1$,

$$\partial m/\partial p_j = x_j - x_s \qquad \text{(B.1}a\text{)}$$

$$\partial m_k/\partial p_j = (x_j - m)^k - k x_j m_{k-1} - [(x_s - m)^k - k x_s m_{k-1}]. \qquad \text{(B.1}b\text{)}$$

Applying equation 3.5 we find that for moments at this locus the elements of the matrix G transform to

$$2g_{k,j} = m_{k+j} - m_k m_j - k m_{j+1} m_{k-1} - j m_{j-1} m_{k+1} + kj m_2 m_{j-1} m_{k-1} \qquad \text{(B.2)}$$

for $k, j > 1$,

$2g_{k,1} = m_{k+1} - k m_2 m_{k-1}$ for $k > 1$, and

$2g_{1,1} = m_2$.

Here $g_{k,j}$ is the element of G corresponding to moments k and j; it is the transformation of equation

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

3.2. Because we assume linkage equilibrium, $g_{k,j} \equiv 0$ for moments of allelic effects at different loci.

The evolution of a locus with s alleles having different effects can be completely described in terms of the evolution of the first $s-1$ moments of allelic effects. This follows because m and m_k for $k = 2, \ldots, s-1$ uniquely determine the first $s-1$ non-central moments, and so yield a system of $s-1$ linear equations for the allele frequencies. (The Vandermonde determinant ensures that this system has a unique solution if all alleles have different effects, see pp. 12–13 of Franklin, 1968.) Although the dynamics of the first $s-1$ moments depend on higher-order moments, the higher moments can be derived from the first $s-1$ by following the inversion recipe given above. From equation B.2, it follows that the upper left corner of the matrix $2G_i$, which governs the dynamics for the moments of locus i, reads:

$$\begin{bmatrix} m_{2i} & m_{3i} & (m_{4i}-3m_{2i}^2) & \cdots \\ m_{3i} & (m_{4i}-m_{2i}^2) & (m_{5i}-4m_{3i}m_{2i}) & \cdots \\ (m_{4i}-3m_{2i}^2) & (m_{5i}-4m_{3i}m_{2i}) & (m_{6i}-m_{3i}^2-6m_{2i}m_{4i}+9m_{2i}^3) & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \quad \text{(B.3)}$$

The complete matrix, G, has one of these matrices for each locus down its diagonal.

Let Z denote a phenotype, let $M = E(Z)$ denote the mean phenotype, and let $M_k = E[(Z-M)^k]$. Our assumptions that all alleles and loci contribute additively to the phenotype, that all genotypes experience independent Gaussian environmental effects with mean 0 and variance V_e, and that all loci are in complete linkage equilibrium imply that the phenotypic moments satisfy

$$M = 2\Sigma_i m_i,$$

$$M_2 = 2\Sigma_i m_{2i} + V_e \equiv V_g + V_e,$$

$$M_3 = 2\Sigma_i m_{3i}, \quad \text{and}$$

$$M_4 = 2\Sigma_i(m_{4i} - 3m_{2i}^2) + 3M_2^2.$$

Thus the dynamic equations 5.1 for the first three moments follow from adding the terms in the first three rows of equation B.3 across loci and noting that

$$\partial(\ln \overline{W})/\partial m_i = 2\,\partial(\ln \overline{W})/\partial M \quad \text{and}$$

$$\partial(\ln \overline{W})/\partial m_{ji} = 2\,\partial(\ln \overline{W})/\partial M_j \quad \text{for } j = 2, 3.$$

(Formulae for fourth and higher moments are slightly more complicated.) Non-additive genetic effects can be incorporated easily if there is an explicit, differentiable function that relates allelic effects to the overall phenotype.

(ii) *Extension to a continuum of alleles*

If linkage disequilibrium is negligible, any polygenic system can be described by the distribution of allelic effects at each locus, $f_i(x_i)$. Consider a randomly mating diploid, and denote the contributions to the phenotype from two alleles at locus i by x_i and x_i^*. First, consider the change in f_i due to selection. If selection is weak enough for changes to be approximately continuous,

$$df_i(x_i)/dt \simeq [W(x_i)f_i(x_i)/\overline{W}] - f_i(x_i). \quad \text{(B.4)}$$

Here $W(x_i)$ is the average fitness of individuals with an allele of effect x_i at locus i, the average being taken over genotypes at other loci, and over the other allele at this locus. To produce an analogue of $\partial(\ln \overline{W})/\partial p_i$ in equation 3.4, we need a concept of differentiation of a functional $S(f)$ with respect to changes in the function f at a specific point, say x_0. We define the functional derivative by (see Gardiner, 1983, section 8.1.1):

$$\partial S(f)/\partial f(x_0) \equiv \lim_{\varepsilon \to 0} \{[S(f(x) + \varepsilon\delta(x - x_0)) - S(f(x))]/\varepsilon\}. \quad \text{(B.5)}$$

where $\delta(x)$ is the Dirac delta function. Using this convention, the gradient of log mean fitness with respect to changes in $f_i(x_i)$ is

$$\partial(\ln \overline{W})/\partial f_i(x_i) = 2W(x_i)f_i(x_i)/\overline{W}. \quad \text{(B.6)}$$

Equations B.4 and B.6 can be related to each other through the generalized indicator covariance matrix: $g(x_i, x_i^*) = f_i(x_i)(\delta(x_i - x_i^*) - f_i(x_i^*))/2$

$$g(x_i, x_j^*) = 0 \quad (i \neq j). \quad \text{(B.7)}$$

This G is now a continuous function of the variables x_i, x_j^*, rather than a discrete matrix.

In this setting, the unpleasant derivatives given by equation B.1 are replaced by

$$\partial m_k/\partial f(x) = (x-m)^k - kxm_{k-1} \quad \text{for} \quad k > 1, \text{ and}$$

$$\partial m/\partial f(x) = x \quad \text{(B.8)}$$

(where we have suppressed the subscript indicating the locus). Proceeding as above leads again to the G matrix for the moments given by equations B.3 and 5.1. This derivation includes that given in Section (i) above, which corresponds to the special case where the distribution $f(x)$ is discrete.

The effects of drift can be found by considering the variance in numbers of genes with allelic effect in the range between x and $x + \delta x$ produced by sampling from the distribution $f(x)$. In the limit of small δx, the binomial sampling variance tends to $g_{x,x^*}/N$, and so is consistent with equation 3.7.

Appendix C

The relative magnitudes of higher-order selection coefficients

If many loci are involved, one expects that the phenotypic distribution will be approximately nor-

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

mal, so that the mean fitness will depend only on the phenotypic mean and variance. To quantify this argument, it is easiest to consider the case in which all loci have identical moments. (However, the argument does not depend on the loci being identical; it suffices that the law of large numbers applies.) Suppose one rescales the allelic effects at each locus relative to the initial standard deviation of the distribution of additive genetic effects, i.e. $\sqrt{V_{g,0}} = \sqrt{2nm_{2,0}}$. Let m_i for $i = 2, 3, \ldots$ denote the central moments at individual loci before rescaling. Then the matrix G governing the dynamics of the scaled variables $\overline{Z}$ and V becomes dimensionless, and the i,jth element is proportional to $n^{-(i+j-2)/2}$:

$$\begin{bmatrix} d\overline{Z}/dt \\ dV/dt \end{bmatrix} \simeq \begin{bmatrix} m_2/m_{2,0} & 2nm_3/(2nm_{2,0})^{3/2} \\ 2nm_3/(2nm_{2,0})^{3/2} & (m_4 - m_2^2)/(2nm_{2,0}^2) \end{bmatrix} \begin{bmatrix} \partial(\ln \overline{W})/\partial\overline{Z} \\ \partial(\ln \overline{W})/\partial V \end{bmatrix} \quad \text{(C.1)}$$

When the number of loci is infinite, the scaled mean changes at a rate equal to the product of the genetic variance (which is one initially) and the derivative of log mean fitness with respect to the mean. To leading order in n, all other moments remain fixed (Bulmer, 1980). When the number of loci is large but finite, the change in each moment is dominated by the first term in each row of G, and depends primarily on the selection coefficient on the mean. However, under stabilizing selection, one would expect the mean to evolve rapidly towards the optimum, where $\partial(\ln \overline{W})/\partial\overline{Z} = 0$. The remaining selection coefficients should be of the same order. When the mean is at equilibrium under stabilizing selection, the first row in G gives

$$(m_2/m_{2,0})\,\partial(\ln \overline{W})/\partial\overline{Z} = -2nm_3/(2nm_{2,0})^{\frac{3}{2}}\,\partial(\ln \overline{W})/\partial V + O(n^{-1}). \quad \text{(C.2)}$$

Thus the first two terms in each row should be of the same order, and the first two selection coefficients must be considered. For example, the change in the variance due to selection is of the order of $1/n$:

$$dV/dt = [-m_3^2/(2nm_2m_{2,0}^2) + (m_4 - m_2^2)/(2nm_{2,0}^2)]\,\partial(\ln \overline{W})/\partial V + O(n^{-\frac{3}{2}}). \quad \text{(C.3)}$$

Of course, stabilizing selection alone would eventually eliminate genetic variance. Under such circumstances, the mean would move rapidly to the vicinity of the optimum (in $t \sim O(1)$ generations), and the mean and variance would then evolve slowly towards a state of complete fixation ($V \to 0$ in $t \sim O(n)$ generations) (see section 7). Equation C.3 describes the latter stage.

If stabilizing selection is balanced by mutation, this argument requires that the mutation pressure on the mean, relative to the phenotypic standard deviation, be small ($O(1/\sqrt{n})$). This is plausible. Suppose that the rate of change of the mean, variance, ... caused by mutation at the ith locus is $\mu_{1i}, \mu_{2i}, \ldots$. The net change in the phenotypic mean, variance... is therefore $2\Sigma_i \mu_{1i}, 2\Sigma_i \mu_{2i}, \ldots$. Now the pressure on the mean may be positive or negative, and is likely to vary independently across loci; however, the variance must increase as a result of mutation. Hence, before rescaling, the expected mutation pressure on the mean is proportional to $\sqrt{n}$, and the mutation pressure on the variance is proportional to n. The effect of mutation on the mean should therefore be $\sim 1/\sqrt{n}$, relative to the variance at a mutation–selection equilibrium. This implies that at equilibrium the selection coefficient on the mean should be of the order of $1/\sqrt{n}$ relative to the higher-order selection coefficients. Therefore, only the first two terms in each row of G, and only the selection coefficients on the mean and the variance, need be considered. However, this argument suggests that the mutation pressure on the mean, though small, may have a significant effect on the variance through the coupling term, M_3, in G. In sections 6 and 7 above this effect was ignored. It requires separate analysis.

To summarize: we expect that when many loci are involved, the selection coefficient on the mean will dominate under directional selection, and the selection coefficients on the mean and the variance will dominate under stabilizing selection.

References

Adler, R., Bazin, M. & Schiffer, M. (1975). *Introduction to General Relativity*. New York: McGraw-Hill.

Akin, E. (1979). *The Geometry of Population Genetics*. Lecture Notes in Biomathematics, 31. Berlin: Springer-Verlag.

Antonelli, P. L. & Strobeck, C. (1977). The geometry of random drift. I. Stochastic distance and diffusion. *Advances in Applied Probability* **9**, 238–249.

Barton, N. H. (1986). The maintenance of polygenic variability by a balance between mutation and stabilizing selection. *Genetical Research* **47**, 209–216.

Barton, N. H. (1987). The effect of linkage and density-dependent regulation on gene flow. *Heredity* **57**, 415–426.

Barton, N. H. & Rouhani, S. (1987). The frequency of shifts between alternative equilibria. *Journal of Theoretical Biology* (In the Press).

Bhattacharya, A. (1946). On a measure of divergence between two multinomial populations. *Sankhya* **7**, 401–406.

Bulmer, M. G. (1972). The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research* **19**, 17–25.

Bulmer, M. G. (1980). *The Mathematical Theory of Quantitative Genetics*. Oxford: Clarendon Press.

Cavalli-Sforza, L. L. & Edwards, A. W. F. (1967). Phylogenetic analysis: models and estimation procedures. *Evolution* **21**, 550–570.

https://doi.org/10.1017/S0016672300026951 Published online by Cambridge University Press

Chakraborty, R. & Nei, M. (1982). Genetic differentiation of quantitative characters between populations or species. I. Mutation and random genetic drift. *Genetical Research* **39**, 303–314.

Crow, J. F. & Kimura, M. (1964). The theory of genetic loads. *Proceedings of the XIth International Congress of Genetics* **2**, 495–505.

Ewens, W. J. (1964). The pseudo-transient distribution and its uses in genetics. *Journal of Applied Probability* **1**, 141–156.

Ewens, W. J. & Thomson, G. (1977). Properties of equilibria in multilocus genetic systems. *Genetics* **87**, 807–819.

Falconer, D. S. (1981). *Introduction to Quantitative Genetics*, 2nd edn. London: Longmans.

Felsenstein, J. (1979). Excursions along the interface between disruptive and stabilizing selection. *Genetics* **93**, 773–795.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Franklin, J. N. (1968). *Matrix Theory*. Englewood Cliffs, N.J.: Prentice-Hall.

Gardiner, C. W. (1983). *Handbook of Stochastic Methods*. Berlin: Springer-Verlag.

Gillespie, J. H. (1984). Pleiotropic overdominance and the maintenance of genetic variation in polygenic traits. *Genetics* **107**, 321–330.

Graham, R. (1977). Covariant formulation of non-equilibrium statistical thermodynamics. *Zeitschrift für Physik B* **26**, 397–405.

Heuch, I. (1975). The relationship between separation time and genetic distance based on angular transformations of gene frequencies. *Biometrics* **31**, 685–700.

Kimura, M. (1965). A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences, U.S.A.* **54**, 731–736.

Kingman, J. F. C. (1978). A simple model for the balance between selection and mutation. *Journal of Applied Probability* **15**, 1–12.

Kirkpatrick, M. (1982). Quantum evolution and punctuated equilibria in continuous genetic characters. *American Naturalist* **119**, 833–848.

Kirkpatrick, M. (1985). Evolution of female choice and male parental investment in polygynous species: the demise of the 'sexy son'. *American Naturalist* **125**, 788–810.

Lande, R. (1975). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research* **26**, 221–235.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334.

Lande, R. (1977). The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* **86**, 485–498.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. *Evolution* **34**, 402–416.

Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**, 203–215.

Lande, R. (1981). Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 3721–3725.

Lande, R. (1985). Expected time for random genetic drift of a population between stable phenotypic states. *Proceedings of the National Academy of Sciences, U.S.A.* **82**, 7641–7645.

Lande, R. & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226.

Latter, B. D. H. (1960). Natural selection for an intermediate optimum. *Australian Journal of Biological Sciences* **13**, 30–35.

Latter, B. D. H. (1970). Selection in finite populations with multiple alleles. II. Centripetal selection, mutation, and isoallelic variation. *Genetics* **66**, 165–186.

Lynch, M. & Hill, W. G. (1986). Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935.

Price, G. R. (1970). Selection and covariance. *Nature* **227**, 520–521.

Schluter, D. (1983). Morphological and phylogenetic relations among the Darwin's finches. *Evolution* **38**, 921–930.

Slatkin, M. (1980). Ecological character displacement. *Ecology* **61**, 163–177.

Slatkin, M. (1987). Heritable variation and heterozygosity under a balance between mutations and stabilizing selection (unpublished).

Turelli, M. (1984). Heritable genetic variation via mutation–selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology* **25**, 138–193.

Turelli, M. (1985). Effects of pleiotropy on predictions concerning mutation–selection balance for polygenic traits. *Genetics* **111**, 165–195.

Turelli, M. (1986). Gaussian versus non-Gaussian genetic analyses of polygenic mutation–selection balance. In *Evolutionary Processes and Theory* (ed. S. Karlin and E. Nevo), pp. 607–628. New York: Academic Press.

Watterson, G. A. (1977). Heterosis or neutrality? *Genetics* **85**, 753–788.

Wright, S. (1935). Evolution in populations in approximate equilibrium. *Journal of Genetics* **30**, 257–266.

Wright, S. (1941). On the probability of fixation of reciprocal translocations. *American Naturalist* **74**, 513–522.

Wright, S. (1949). Adaptation and selection. In *Genetics, Paleontology, and Evolution* (ed. G. G. Simpson, G. L. Jepson and E. Mayr), pp. 365–389. Princeton: Princeton University Press.

Wright, S. (1965) Factor interaction and linkage in evolution. *Proceedings of the Royal Society* B **162**, 80–104.

Wright, S. (1969). *Evolution and the Genetics of Populations*, vol. 2. *The Theory of Gene Frequencies*. Chicago: University of Chicago Press.