lution, 43(6), 1989, pp. 1209–1222

VISUALIZING MULTIVARIATE SELECTION[1]

PATRICK C. PHILLIPS AND STEVAN J. ARNOLD
Committee on Evolutionary Biology and Department of Ecology and Evolution, University of Chicago, 915 E. 57th Street, Chicago, IL 60637

Abstract.—Recent developments in quantitative-genetic theory have shown that natural selection can be viewed as the multivariate relationship between fitness and phenotype. This relationship can be described by a multidimensional surface depicting fitness as a function of phenotypic traits. We examine the connection between this surface and the coefficients of phenotypic selection that can be estimated by multiple regression and show how the interpretation of multivariate selection can be facilitated through the use of the method of canonical analysis. The results from this analysis can be used to visualize the surface implied by a set of selection coefficients. Such a visualization provides a compact summary of selection coefficients, can aid in the comparison of selection surfaces, and can help generate testable hypotheses as to the adaptive significance of the traits under study. Further, we discuss traditional definitions of directional, stabilizing, and disruptive selection and conclude that selection may be more usefully classified into two general modes, directional and nonlinear selection, with stabilizing and disruptive selection as special cases of nonlinear selection.

Received March 22, 1988. Accepted March 20, 1989

Recent developments in quantitative-genetic theory have yielded a coherent view of how continuously varying phenotypic traits evolve within populations (e.g., Lande, 1979, 1980, 1988). Implicit in this approach is a multivariate characterization of natural selection acting on many traits simultaneously (Thompson, 1977; Lande and Arnold, 1983). This multidimensional view of selection contrasts with standard textbook accounts of directional and stabilizing selection, which usually take a univariate approach (e.g., Grant, 1977; Ayala and Valentine, 1979; Cavalli-Sforza and Feldman, 1981; but see Schmalhausen [1949] and Lerner [1954]). While the empirical literature shows an increasing tendency to consider multivariate selection, usually only directional selection is analyzed (e.g., Price et al., 1984; Berenbaum et al., 1986; but see Price and Boag [1987]). Further, the importance of multivariate stabilizing selection in visualizations of selection has not been emphasized in recent monographs on natural selection (Manly, 1985; Endler, 1986).

The univariate view of selection may prevail because it is difficult to conceptualize selection acting on many traits at once. For example, in the univariate case, it is often convenient to picture selection in terms of a fitness function acting on a continuously varying trait (Fig. 1), whereas a multivariate approach rapidly exhausts our capacity for visualization. Despite this difficulty, it is well worth the effort to view selection in a multidimensional context, for it is only in this way that we can understand how selection acts on more than one trait at a time and, in particular, how selection affects and is affected by the correlation between characters. Our aim in this paper is, therefore, to review some recent concepts of multivariate selection and to show how these concepts can be visualized.

In the sections that follow, we briefly review terminology and concepts of phenotypic selection in order to point out connections with genetic change and with multivariate visualization of selection. We then introduce canonical analysis of selection surfaces, a technique that helps in the interpretation of systems of selection coefficients.

Modes of Selection

Natural selection has traditionally been classified into three modes: directional, stabilizing, and disruptive (e.g., Kimura, 1983 pp. 119–121; Endler, 1986 pp. 16–21). These univariate modes of selection are usually illustrated by a fitness function described by a line, peak, or valley (Fig. 1a, b). Stabilizing (disruptive) selection is thus characterized by the existence of an intermediate opti-

[1] This paper is dedicated to the memory of Sewall Wright (1889–1988).

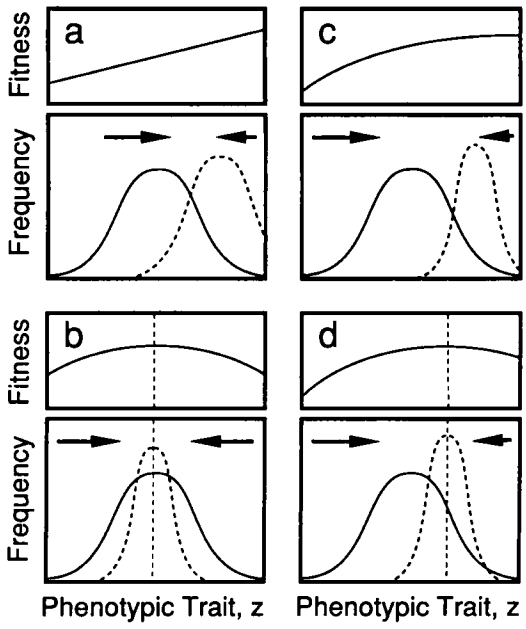

FIG. 1. Standard univariate views of selection acting on a single trait, showing effects on trait distributions within a single generation. Solid, bell-shaped outlines indicate the trait distribution before selection; stippled outlines show the trait distribution after selection. Selection functions are shown above frequency distributions: a) Directional selection under all definitions, b) stabilizing selection under all definitions; c) traditionally defined as only directional selection, defined as a combination of directional selection and stabilizing selection by Lande and Arnold (1983), and defined as a combination of directional and nonlinear selection in this paper, d) a combination of directional and stabilizing selection under all definitions.

mum (minimum) within the range of phenotypic expression. Lande and Arnold (1983) proposed a slightly more general definition of stabilizing and disruptive selection: curvature in the relationship between fitness and the trait. These two definitions come into conflict, however, when the fitness function is curved but monotonic increasing or decreasing (Fig. 1c). This could occur, for example, if fitness approaches some asymptotic value with increasing trait values, or if the selective optimum is simply not within the range of the current phenotypic distribution. Such a fitness function would traditionally be defined as purely directional selection (e.g., Kimura, 1983; Manly, 1985; Endler, 1986; Mitchell-Olds and Shaw, 1987; Schluter, 1988), whereas Lande and Arnold (1983) would classify this as a combination of directional and stabilizing selection. Thus, there is potential for confusion because of these two somewhat distinct uses of terminology (see Schluter [1988] for an example).

Modes of selection are probably most usefully defined in terms of the changes in the phenotypic distribution that are caused by selection (Lande and Arnold, 1983; Endler, 1986 p. 17). These changes can be roughly divided into 1) change in the mean and 2) changes in all other moments of the phenotypic distribution. The change in the mean of a phenotypic trait within a generation is the implied meaning of "directional selection," and therefore, the relationship between fitness and the trait that describes the change in mean provides the best definition of directional selection. In the multivariate case, this relationship is given by the partial linear regression of fitness on a set of characters (see below; Lande and Arnold, 1983). Thus, directional selection is seen to be a fundamentally linear process and has, in fact, been called linear selection by Simpson (1953) and Spiess (1977).

The connection between directional selection and the linear relationship between fitness and a trait suggests that the other modes of selection be defined in terms of the nonlinear relationships between fitness and traits that cause changes in the higher moments of the phenotypic distribution. Indeed, the presence of stabilizing selection has often been inferred by the contraction of variance (reviewed by Endler [1986]). However, directional selection by itself can also cause a change in variance (Lande and Arnold, 1983), so only changes in the higher moments beyond those caused by a change in the mean should be considered to be indicative of the presence of stabilizing or disruptive selection. It is this differentiation between the change in variance caused by directional selection and the change in variance attributable to stabilizing selection that led Lande and Arnold (1983) to define stabilizing and disruptive selection in terms of curvature in the fitness function. For historical reasons and to avoid confusion, however, it seems desirable to retain the traditional definitions of stabilizing and disruptive selection. In particular, stabilizing selection was originally proposed as a description of a population in selective equi-

librium (Schmalhausen, 1949; Waddington, 1957), and the existence of a selective equilibrium implies a peak in the fitness surface (e.g., Fig. 1b; Wright, 1969, 1977; Lande, 1976, 1979).

It is important to realize, however, that the traditional definitions of stabilizing and disruptive selection are typological and qualitative. For example, in some instances, contraction of variance beyond that caused by directional selection will be indicative of stabilizing selection (Fig. 1d), whereas in other instances it is not (Fig. 1c). Further, as a population moves toward a fitness peak, selection changes from purely directional (Fig. 1c) to a combination of directional and stabilizing selection (Fig. 1d). The point at which stabilizing selection begins can only be defined arbitrarily. This confusion over the meaning of stabilizing selection has not been a problem in theoretical work, because the fitness function is usually defined independently of the location of the population mean. In empirical studies of selection, however, there is a potential for confusion between labels attached to selection coefficients and the traditional modes of selection. Much of the confusion can be eliminated by recognizing a new, general category of selection. We propose that *nonlinear selection* be used to describe selection that causes a change in the second or higher moments of the phenotypic distribution beyond those caused by directional selection. Under this system, then, there are two fundamental modes of selection: directional (linear) and nonlinear. In the univariate case, nonlinear selection can be divided into convex selection (with the fitness function bent downward) and concave selection (the fitness function bent upward). Stabilizing and disruptive selection are thus special cases of convex and concave selection in which an inflection point on the fitness surface occurs in the neighborhood of the phenotypic mean (e.g., within the range of observed phenotypic variation). We leave unresolved the problem of usefully defining the term, "in the neighborhood of the phenotypic mean" (see Mitchell-Olds and Shaw, 1987).

Because of the assumption of multivariate normality, which we will use throughout our discussion, the change in the mean and variance of the phenotypic distribution can be predicted by the linear and quadratic elements of the fitness function (Lande and Arnold, 1983). Therefore, in this paper the only coefficients of nonlinear selection that we will consider will be coefficients of quadratic selection.

Multivariate Approaches to Selection

A simple bivariate association between individual fitness and a trait (as in Fig. 1) is commonly used in the empirical literature as a test for directional or stabilizing selection (reviewed by Endler [1986 Ch. 5]). A limitation of such univariate measures of selection is that the association may be due to common correlation with some other trait (Pearson, 1903; Robertson, 1956; Waddington, 1957; Falconer, 1981 Ch. 19). For example, if we are measuring selection on trait A and if another trait (B) is correlated with A, selection acting only on B can make it appear that there is a causal association between A and fitness, even if A is itself selectively neutral. If both traits have been measured, however, one can attempt to tease apart the direct and indirect effects of selection by computing partial correlations or partial regressions. While even the multivariate approach is plagued to some degree by the problem of unmeasured characters, it has the advantage that it can correct for characters that are measured (Lande and Arnold, 1983). Some researchers have used this multivariate approach with a justification rooted only in statistical grounds (e.g., McGregor et al., 1981). Lande and Arnold (1983) showed that a partial-regression approach to selection also makes sense in terms of evolutionary theory, because the regression coefficients are also the selection coefficients that appear in dynamical equations for phenotypic evolution and that define a multidimensional surface which describes the relationship between the traits and fitness.

Selection Coefficients.—In the particular system of selection coefficients that we will use, each coefficient has three different meanings: 1) a dynamic meaning, describing the change in the trait distribution caused by selection; 2) a statistical meaning, describing the relationship between a phenotypic trait and fitness; and 3) a geometrical meaning, describing a surface of individual

TABLE 1. A summary of selection coefficients and their meanings.

Selection coefficient	Symbol	Dynamic interpretation	Statistical meaning	Geometrical interpretation
Directional selection differential	s_i	shift in mean due to direct and indirect effects of directional selection	covariance between relative fitness and character	—
Directional selection gradient	β_i	strength of the direct force of directional selection on character z_i	partial regression of relative fitness on a character, holding all other characters constant[b]	direction of steepest uphill slope from the population mean on the selection surface[a]
Quadratic selection differential	C_{ij}	change in variance ($i = j$) of character z_i or covariance ($i \neq j$) between characters z_i and z_j due to the direct and indirect effects of quadratic selection, independent of the influence of directional selection	covariance between relative fitness and pairwise products of character deviations from the mean	—
Quadratic selection gradient	γ_{ij}	strength of the direct effects of quadratic selection: when $i = j$, γ_{ii} indicates whether convex (negative γ) or concave selection (positive γ) is acting on trait z_i; when $i \neq j$, γ_{ij} indicates the direct effects of correlational selection	partial regression of relative fitness on pairwise products of character deviations from the mean, holding other characters constant[b]	curvature and orientation of the selection surface[a]

[a] Assumes multivariate normality of the phenotypic distribution.
[b] Assumes multivariate normality only if coefficients are not estimated separately.

fitness as a function of trait values (Lande and Arnold, 1983). The coefficients and their meanings are listed in Table 1. The coefficients are of two basic types. The *selection differentials* most closely resemble the common-sense measures of selection, for they represent the changes in mean, variance, and covariance that are induced by selection. Thus, for example, the directional selection differential (shift in mean due to selection) has been used for many years in numerous applications in quantitative genetics (Lush, 1945; Falconer, 1981 Ch. 11). Similarly, the quadratic selection differential measures changes in trait variance and covariance caused by selection. The fundamental limitation of selection differentials as measures of selection is that they reflect the impact of selection on correlated characters as well as on the trait in question. Thus, the selection differentials show the results of both the direct effects of selection on a given trait and the indirect effects of selection on correlated traits. However, using the fact that the selection differentials are equivalent to covariances between fitness and the traits (Table 1), one can compute multivariate measures of selection that correct for correlations among the measured traits and, therefore, only estimate the direct effects of selection (Lande and Arnold, 1983). These multivariate selection coefficients are partial regressions of fitness on the traits. They have been called *selection gradients* (Table 1).

Because they measure only the direct effects of selection on a trait, the directional selection gradients can be used in dynamic equations to predict the evolution of the average phenotype in the population. Thus, if we wish to predict how much the mean of a set of traits will change from one gen-

eration to the next due to the deterministic effects of selection, and if the traits have a multivariate normal distribution, we can use the equation

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta} \tag{1}$$

where $\Delta\bar{\mathbf{z}}$ denotes a column vector of changes in phenotypic means, **G** denotes the additive genetic variance-covariance matrix for the traits, and $\boldsymbol{\beta}$ is the directional selection gradient vector (Lande, 1979). The quadratic selection gradients play an analogous role in a deterministic equation for the change in the genetic matrix **G**. Thus, the change in **G** due to selection within a generation (assuming multivariate normality of the traits) is

$$\Delta\mathbf{G} = \mathbf{G}(\boldsymbol{\gamma} - \boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}})\mathbf{G} \tag{2}$$

where $\boldsymbol{\gamma}$ is the quadratic selection gradient matrix and T denotes matrix transposition (after Lande [1980] and Lande and Arnold [1983]). How these within-generation changes in **G** are transmitted across generations is somewhat controversial (see Lande, 1980, 1984; Turelli, 1984, 1985).

Fitness Surfaces

Another important meaning of the selection gradients is that they describe a selection surface. Three different types of fitness surfaces are recognized in the evolutionary literature: the individual selection surface, the best quadratic approximation to the individual selection surface, and the adaptive landscape. Selection gradients describe the second type of surface and serve as a link between the other two.

The Individual Selection Surface.—The individual selection surface represents the association between the expected fitness of an individual and its phenotypic value(s) for various traits (Fig. 1). From an ecological perspective, the individual selection surface describes the fitness consequences of phenotypes interacting with their environment. In the absence of frequency-dependent selection, then, the individual selection surface can be thought of as a feature of the environment that is independent of the distribution of the phenotypes (Schluter, 1988). The individual selection surface could be of a simple, smooth shape (e.g., quadratic), or it might be highly irregular. The "true" surface is therefore probably best estimated using techniques that do not make assumptions about the functional form of the surface (see Schluter [1988] for some useful techniques that complement those presented here). A quadratic approximation of the individual selection surface is very useful, however, because it yields estimates of the selection gradients.

The Best Quadratic Approximation to the Individual Selection Surface.—As discussed above, the selection gradients are actually coefficients of a quadratic regression of fitness on the measured traits. This regression provides the best quadratic approximation of the individual selection surface. The gradients can be estimated using standard multiple-regression techniques, which, after standardizing the trait means to zero, results in the equation

$$w = \alpha + \sum_{i=1}^{n} \beta_i z_i + \sum_{i=1}^{n} \frac{1}{2}\gamma_{ii} z_i^2 + \sum_{i<d}^{n}\sum^{n} \gamma_{ij} z_i z_j + \epsilon \tag{3}$$

where w is relative fitness (absolute fitness divided by mean absolute fitness), α is a constant, β_i is the directional selection gradient for trait z_i, γ_{ii} is the quadratic selection gradient for trait z_i (indicating concave or convex selection), γ_{ij} is the quadratic selection gradient for traits z_i and z_j (indicating correlational selection), and ϵ is an error term (after Lande and Arnold [1983 eq. 16]). For a given phenotypic distribution, however, many different individual selection surfaces can yield the same selection gradients, because the directional and quadratic selection gradients represent the *average* slope and curvature of the individual surface (Lande and Arnold, 1983). Thus, the selection gradients can also be expressed as the weighted partial derivatives

$$\boldsymbol{\beta} = \int p(\mathbf{z}) \frac{\partial w(\mathbf{z})}{\partial \mathbf{z}}\, d\mathbf{z}, \tag{4}$$

and

$$\boldsymbol{\gamma} = \int p(\mathbf{z}) \frac{\partial^2 w(\mathbf{z})}{\partial \mathbf{z}^2}\, d\mathbf{z}, \tag{5}$$

where $\partial/\partial\mathbf{z} \equiv (\partial/\partial z_1, \ldots, \partial/\partial z_n)^{\mathrm{T}}$ is the gradient (slope) operator, where

$$\partial^2/\partial \mathbf{z}^2 \begin{bmatrix} \partial^2/\partial z_1^2 & \cdots & \partial^2/\partial z_1 \partial z_n \\ \vdots & \ddots & \vdots \\ \partial^2/\partial z_n \partial z_1 & \cdots & \partial^2/\partial z_n^2 \end{bmatrix}$$

is the curvature operator, and where integration is taken over all phenotypes in the population (cf. Lande and Arnold, 1983 eq. 9 and eq. 14b).

The relationship between the selection gradients and the individual selection surface can be seen by taking the Taylor expansion of the individual selection surface around the mean phenotype in the population. Using the univariate case for notational convenience and assuming normality, (4) and (5) yield

$$\beta = \left[\frac{dw(z)}{dz} + \frac{\sigma^2}{2}\frac{d^3w(z)}{dz^3} + \frac{\sigma^4}{8}\frac{d^5w(z)}{dz^5} + \cdots\right]_{z=\bar{z}} \quad (6)$$

and

$$\gamma = \left[\frac{d^2w(z)}{dz^2} + \frac{\sigma^2}{2}\frac{d^4w(z)}{dz^4} + \frac{\sigma^4}{8}\frac{d^6w(z)}{dz^6} + \cdots\right]_{z=\bar{z}} \quad (7)$$

where σ^2 is the variance of the phenotypic trait and $[d^iw(z)/dz^i]_{z=\bar{z}}$ is the ith derivative of the individual selection surface evaluated at the phenotypic mean. These relationships show that the slope and curvature of the individual selection surface at the population mean will be the same as β and γ only if the actual surface is quadratic, for only then will the higher order terms in (6) and (7) be zero. Otherwise, the relationship between β and γ and the individual selection surface will depend on both the shape of the surface and the phenotypic distribution, and therefore, Equation (3) may provide a poor approximation to the true individual selection surface.

Finally, it should be noted that, while the individual selection surface represents the relationship between an *individual* and fitness, selection gradients are necessarily attributes of the *population* under study, since they are used to describe changes in population characteristics [i.e., means and variances; Table 1; Eqs. (1) and (2)]. Hence, only the average properties of the individual selection surface are necessary to describe the evolutionary dynamics of the population, although the particular form of the individual selection surface may be of interest for a functional interpretation of fitness differences between individuals.

Adaptive Landscapes.—Before continuing, it should be noted that the selection surfaces discussed above are qualitatively different from the third type of fitness surface, the adaptive landscape (Wright, 1932, 1977; Simpson, 1953). Wright used a landscape in which the vertical dimension was the mean fitness in the population and the horizontal dimensions were also population attributes (gene frequencies). Simpson suggested that this idea might be applicable to phenotypic characters, and the relationship was made precise by Lande (1976, 1979). Such adaptive landscapes have the useful property that the evolving population will tend to move uphill on the surface and, if fitnesses are constant, will eventually equilibrate on a local peak (Lande, 1976, 1979; Wright, 1977). The selection surface referred to in this paper is not such an adaptive landscape but is, instead, a surface of individual fitness as a function of individual phenotypic trait values.

The relationship between the individual selection surface and the adaptive landscape can be interpreted, however, using the best quadratic approximation to the individual selection surface. The selection gradients, β and γ, are, respectively, the average slope and average curvature of the individual selection surface weighted by the phenotype distribution [Eqs. (4) and (5)]. They are related to the slope and curvature of the adaptive landscape evaluated at the multivariate mean of the trait distribution. In particular, β is also the slope of the adaptive landscape (cf. Lande, 1979 eq. 6), while the curvature of the landscape is $\gamma - \beta\beta^T$ if the phenotypic trait distribution is multivariate normal [Lande, pers. comm.; note its role in Eq. (2)]. Thus, when directional selection is absent ($\beta = \mathbf{0}$), the adaptive landscape and the best quadratic fit to the individual selection surface [Eq. (3)] have the same curvature at the optimum (e.g., Fig. 2a). When there is directional selection, however, the curvature of the adaptive landscape at the population mean is always more negative (i.e.,

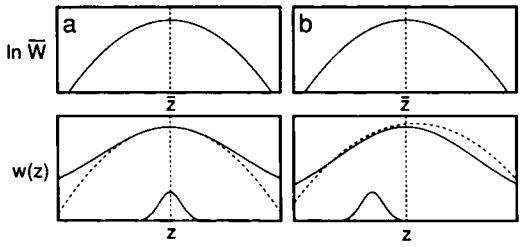

FIG. 2. Correspondence between the adaptive landscape, the actual individual selection surface, and the fitted quadratic selection surface. The top two graphs show the adaptive landscape i.e., the natural logarithm of mean absolute fitness as a function of mean phenotype; the bottom two graphs show the individual selection surface (solid curves), i.e., individual relative fitness as a function of phenotype. The fitted selection surface (dashed curves in the bottom graphs) is the best quadratic approximation of the individual selection surface. The phenotypic trait distribution is shown as a bell-shaped curve. The vertical dashed line shows the selective optimum. The graphs were generated assuming a Gaussian individual fitness surface and a normal phenotypic distribution with standard deviation one tenth the width of the fitness surface. a) Nonlinear (stabilizing) selection but no directional selection; note that the curvatures of the adaptive landscape and the fitted surface are the same. b) Nonlinear and directional selection; note that, at the mean, the curvature of the adaptive landscape is greater than that of the fitted surface.

more curved when convex selection is operating) than the average curvature of the individual selection surface at that point (e.g., Fig. 2b).

Interpreting the Selection Gradients.— Given a surface described by (3), the coefficients from the quadratic regression describe the slope, curvature, and orientation of the selection surface (Table 1). The vector of directional selection gradients ($\boldsymbol{\beta}$) indicates the direction of steepest uphill slope on the surface from the population mean ($\bar{\mathbf{z}} = \mathbf{0}$). The sign of β_i describes whether trait z_i is under positive or negative directional selection, while its magnitude describes the strength of that selection. Thus, a selection surface that is described only by directional selection gradients will be a tilted plane, as is depicted in Figure 3a, b for selection on two traits.

Despite the ease of their calculation, interpretation of the quadratic selection gradients turns out to be somewhat more complex than that of the directional selection gradients. The diagonal elements of the quadratic selection gradient matrix (γ_{ii}) describe the curvature of the surface along the individual trait axes (z_i). A negative value for γ_{ii} means that the fitness surface is curved downward (convex selection; e.g., Fig. 1b–d), while a positive value of γ_{ii} indicates that the surface is curved upward (concave selection). However, selection can also act on combinations of traits (correlational selection). The pattern of correlational selection

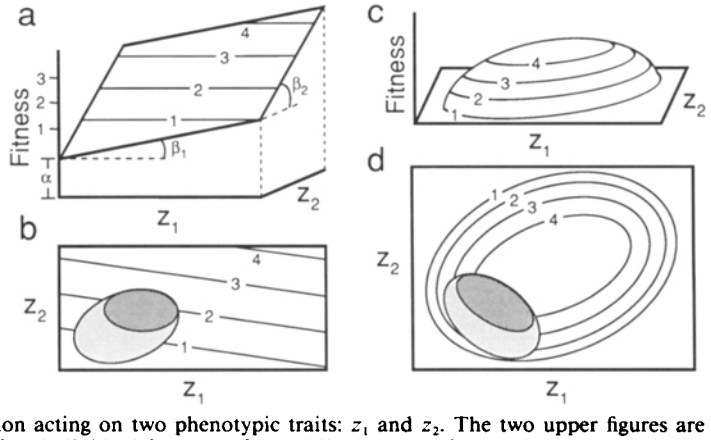

FIG. 3. Selection acting on two phenotypic traits: z_1 and z_2. The two upper figures are three-dimensional representations of an individual fitness surface, while the lower figures show equal-height contour plots of the same surfaces. a) Positive directional selection on both traits. Directional selection gradients (β_1 and β_2) indicate the slope of the surface; α is a constant of elevation (after Box and Draper, 1987 fig. 2.1). b) The effect of bivariate directional selection on the distribution of a population. Stippled ellipses show the population distribution before (light stippling) and after (dark stippling) selection within one generation. Outlines of the ellipse can be thought of as the 95% confidence intervals of the phenotypic distribution. c) Bivariate convex selection with positive correlational selection. Surface represents $\gamma_{22} < \gamma_{11} < 0 < \gamma_{12}$, and $\gamma_{12}^2 < \gamma_{11}\gamma_{22}$ (see text for an explanation of symbols). d) The effect of convex selection on a population distribution.

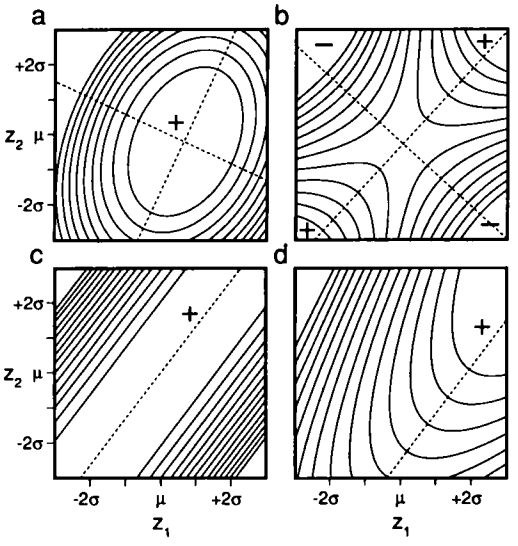

FIG. 4. Hypothetical individual selection surfaces for two phenotypic traits (+ = peak; − = depression). Here, β_1 and β_2 are the directional selection gradients; γ_{11}, γ_{22}, and γ_{12} are the quadratic selection gradients; and λ_1 and λ_2 are the canonical coefficients of the system (eigenvalues of the γ matrix). Dotted lines represent the canonical (principal) axes of the system. For all surfaces, except where otherwise indicated, $\beta_1 = 0.22$, $\beta_2 = -0.08$, $\gamma_{11} = -0.31$, $\gamma_{22} = -0.18$. a) Selection surface with a peak ($\gamma_{12} = 0.07$, $\lambda_1 = -0.15$, $\lambda_2 = -0.34$); b) selection surface with a saddle ($\gamma_{12} = 0.77$, $\lambda_1 = 0.53$, $\lambda_2 = -1.01$); c) selection surface with a ridge ($\beta_1 = 0$, $\beta_2 = 0$, $\gamma_{12} = 0.23$, $\lambda_1 = 0$, $\lambda_2 = -0.48$); d) selection surface with a rising ridge ($\beta_1 = 0.84$, $\gamma_{12} = 0.23$, $\lambda_1 = 0$, $\lambda_2 = -0.48$.

on traits z_i and z_j is revealed in γ_{ij}. A positive value for γ_{ij} indicates that traits z_i and z_j are being selected to become positively correlated; the opposite is true for a negative value. If correlational selection on a set of traits is absent or weak, then the univariate interpretation of the individual γ_{ii} values applies to the multivariate surface as a whole. For instance, with two traits, if both γ_{11} and γ_{22} are negative and γ_{12} is small, the fitness surface will be a hill (Fig. 3c, d). In general, however, when more than one trait is involved, nonzero values of γ_{ij} complicate the interpretation of the multivariate surface implied by the β and γ matrices. For example, all of the surfaces drawn in Figure 4 possess similar selection coefficients, yet the surfaces are quite different. In particular, in Figure 4b both γ_{11} and γ_{22} are negative, leading one naively to expect the surface to be a peak when it is in fact a saddle. Similar difficulties are illustrated in Figure 4c, d. It should be noted that Figure 4a–c all share the same coefficients except for differences in the magnitude of γ_{12} (Fig. 4d is a special case of Fig. 4c and will be discussed below).

Quadratic Surfaces.—In a given dimensional space, the number of distinct types of quadratic surfaces is finite. For instance, in three dimensions (i.e., measuring selection on two traits) the possible surfaces are illustrated in Figure 4: a peak or valley, a saddle, a ridge, and a rising ridge. Quadratic surfaces also have the property that they are symmetrical about a particular set of axes called the major axes (Fig. 4). For a surface constructed by an analysis of selection on n characters, there will be n major axes of the surface. The symmetry of this system suggests a solution to the problem illustrated in Figure 4. If the surface could be interpreted in terms of the major axes, then there would be no confusion as to the form of curvature along a given axis. This interpretation can be achieved by a relatively simple transformation which rotates the original axes to the major axes, thereby removing the influence of the γ_{ij} coefficients.

Canonical Analysis

The problem of interpreting and visualizing quadratic selection surfaces is a subset of the statistical issues involved in response-surface analysis (Box and Draper, 1987). Box and his coworkers have developed many techniques for analyzing surfaces approximated by quadratic regression (i.e., response surfaces; Box and Wilson, 1951; Box 1954; Box and Youle, 1955; Box and Draper, 1987; see Mead and Pike [1975] for some current uses and historical background). Central to this methodology is canonical analysis, which is a method of rewriting a fitted second-degree equation in a form that can be more readily interpreted (Box and Draper, 1987 p. 332). This new interpretation is achieved by a translation and rotation of the coordinates of the multidimensional space so that the new axes are aligned with the major axes of the fitted surface.

It should be noted that the canonical analysis described here is not the same as canonical-correlation or canonical-variate analysis, although they share similar statis-

tical techniques, and are sometimes called by this name. Canonical analysis, as we use it, is a set of general methods for characterizing matrices, of which the methods used in principal-components analysis are a subset.

To proceed with the analysis, note that Equation (3) can be rewritten in matrix form (without the error term) as

$$w = \alpha + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{z} + \tfrac{1}{2}\mathbf{z}^{\mathrm{T}}\boldsymbol{\gamma}\mathbf{z} \qquad (8)$$

where, as before, all traits have been standardized to zero means. Information about the curvature and orientation of the surface is contained in the $\boldsymbol{\gamma}$ matrix. In order to obtain the necessary rotation of axes, note that $\boldsymbol{\gamma} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^{\mathrm{T}}$, where $\mathbf{M}$ is an orthogonal matrix whose columns are the eigenvectors of $\boldsymbol{\gamma}$ normalized to unit length, and $\boldsymbol{\Lambda}$ is a matrix with the eigenvalues of $\boldsymbol{\gamma}$ on its diagonal and zeros everywhere else. From the preceding relationships, it can be seen that the transformation

$$\boldsymbol{\Lambda} = \mathbf{M}^{\mathrm{T}}\boldsymbol{\gamma}\mathbf{M} \qquad (9)$$

results in a new matrix of multivariate curvature ($\boldsymbol{\Lambda}$) in which all off-diagonal elements are zero. Now, writing $\mathbf{y} = \mathbf{M}^{\mathrm{T}}\mathbf{z}$ and $\theta = \mathbf{M}^{\mathrm{T}}\boldsymbol{\beta}$, (8) can be rewritten as

$$w = \alpha + \boldsymbol{\theta}^{\mathrm{T}}\mathbf{y} + \tfrac{1}{2}\mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y}, \qquad (10a)$$

or, in long form, as

$$w = \alpha + \theta_1 y_1 + \ldots + \theta_n y_n + \tfrac{1}{2}(\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2). \qquad (10b)$$

Equation (10) is the "A canonical form" of Box and Draper (1987 p. 333). We will call the diagonal elements of $\boldsymbol{\Lambda}$ (λ_i = the eigenvalues of $\boldsymbol{\gamma}$) the *canonical coefficients* of the system.

A great deal of information can be obtained from Equation (10). The signs of the λ_i determine the type of fitted second-order surface, and their magnitudes describe the curvature of the surface. The θ_i measure the slope of the surface from the original origin ($\bar{\mathbf{z}} = \mathbf{0}$) along the rotated axes indicated by the transformed variables, y_i. Before a precise interpretation of the λ_i can be explained, however, one further transformation is needed.

Because Equation (8) represents a surface, the stationary point, $\mathbf{z}_0$ (a minimum, maximum, or saddle point) on that surface can be found by setting its derivative to zero, which results in

$$\mathbf{z}_0 = -\boldsymbol{\gamma}^{-1}\boldsymbol{\beta}. \qquad (11)$$

By substituting (11) into (8), the fitness value at the stationary point can be found, and this is given by

$$w_0 = \alpha + \tfrac{1}{2}\boldsymbol{\beta}^{\mathrm{T}}\mathbf{z}_0. \qquad (12)$$

The system is then put in "B canonical form" (Box and Draper, 1987 p. 337) by shifting the origin to the stationary point and rotating the axes as above by the transformation, $\mathbf{y} = \mathbf{M}^{\mathrm{T}}(\mathbf{z} - \mathbf{z}_0)$. Under this transformation and using the decomposition of $\boldsymbol{\gamma}$ given above, (8) can be rewritten as

$$w = w_0 + \tfrac{1}{2}\mathbf{y}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{y} \qquad (13a)$$

or

$$w = w_0 + \tfrac{1}{2}(\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2). \qquad (13b)$$

Except for the linear terms, the interpretation of the transformation in form A [Eq. (10)] and form B [Eq. (13)] is the same. Here, the important factors are the sign and magnitude of the canonical coefficients, λ_i. In general, if all the λ_i are negative, then $\mathbf{z}_0$ is a point of maximum fitness (such as in Fig. 4a), indicating that convex selection is operating on all traits and trait combinations. Similarly, if all the λ_i are positive then $\mathbf{z}_0$ is a minimum-fitness point, with concave selection acting on the character distribution. If the signs of the λ_i are different, however, then $\mathbf{z}_0$ is an unstable equilibrium, and the fitness surface will look like a saddle (Fig. 4b). The larger the magnitude of the canonical coefficients ($|\lambda_i|$), the more curved the surface will be in that dimension. Therefore, if some of the λ_i are nearly zero, there will not be much curvature in the surface along the axes associated with these values. Such a situation means that there is a ridge of almost constant fitness along one axis (or set of axes; e.g., Fig. 4c). An eigenvalue of zero indicates that the $\boldsymbol{\gamma}$ matrix is singular and, therefore, that Equation (11) for calculating the stationary point cannot be computed. Canonical form A should always be

used to characterize such a system. This is especially important when directional selection is strong relative to quadratic selection and one or more of the eigenvalues is zero. In this case, a rising ridge will result (e.g., Fig. 4d), with the slope of the ridge equal to θ_i, where $\lambda_i = 0$ [Eq. (10)]. In general, form A should be used whenever the stationary point is beyond the distribution of observed phenotypes, because a translation of the origin to a stationary point in unmeasured space may lead to erroneous interpretations of the fitted surface. Form B can be used when the stationary point is within the phenotypic range and when one wants the relationship between the population mean and the stationary point to be emphasized (e.g., under stabilizing or disruptive selection).

A problem with this technique, as with all transformations, is that the transformed variables must be interpreted in terms of the original variables. However, in this case, the transformation is orthogonal, and the relationship between the original variables and the transformed variables is contained in the eigenvector matrix, **M**. Hence, each column in **M** contains the loadings of original variables on the transformed axes. These loadings indicate the linear combination of traits in the original space (**z**) that create the new traits in the canonical space (**y**) and can therefore be interpreted as in principal-components analysis.

Selection on Two Traits.—When measuring selection on only two traits, some simple rules relating the above results directly to the γ matrix can be formulated. In this two-dimensional case, performing the canonical analysis is simplified, because solving for the eigenvalues of γ only necessitates using the quadratic equation. This allows one to find an explicit relationship between the diagonal and off-diagonal terms in γ. In particular, if γ_{11} and γ_{22} have the same sign, then it can be shown that the relationships $\gamma_{12}^2 < \gamma_{11}\gamma_{22}$, $\gamma_{12}^2 = \gamma_{11}\gamma_{22}$, and $\gamma_{12}^2 > \gamma_{11}\gamma_{22}$ result in the surface being a peak (if both diagonal elements are negative; a valley if both are positive), a ridge, and a saddle, respectively. If the signs of γ_{11} and γ_{22} are different, however, the surface will always be a saddle. Thus, variation in the value of γ_{12}^2 around the point $\gamma_{12}^2 = \gamma_{11}\gamma_{22}$ can lead to qualitatively different types of surfaces (e.g., Fig. 4). Analogous results are given in Kimura (1956) and Wright (1969 pp. 41–42) for multidimensional adaptive landscapes.

Statistical Issues.—We do not wish to address the statistical problems associated with estimating selection gradients in natural populations per se, as many aspects of this subject have been covered elsewhere (Lande and Arnold, 1983; Mitchell-Olds and Shaw, 1987). Canonical analysis is simply a post hoc, orthogonal transformation of the original variables and, as such, introduces no statistical error through the transformation itself. Therefore, if the selection gradients have been poorly estimated, canonical analysis will not improve their accuracy. Errors can be assigned to the canonical coefficients of the estimated selection gradients, however. In general, if the phenotypic distribution is fairly uniform across all characters (i.e., if there is little correlation between characters), then the axes of the surface are freely rotatable, and the errors of the canonical coefficients (λ_i) will be on the same order as the quadratic selection gradients (γ_{ii}) (Box and Draper, 1987 p. 354). Such a distribution is unlikely to be obtained in real populations, however, so the errors of the canonical coefficients will need to be calculated. Unfortunately, because of the difficulties of assigning errors to eigenvalues, no direct transformation of the errors in γ to those in λ is known. The errors can be readily calculated, however, by simply transforming the variables to the canonical space before analysis and then calculating the regression specified in Equations (10) or (13) (E. Simms, unpubl.). The regression coefficients estimated by this procedure will be the canonical coefficients, and their errors can be calculated in the same way that one would calculate an error for any regression coefficient (this procedure is analogous to using principal components as the variables in a regression; Draper and Smith, 1981).

Solutions for the confidence region of the stationary point [Eq. (11)] are given by Box and Hunter (1954; see also Stablein et al. [1983]). We do not see the prediction of selective optima as the primary use of this method, however. In particular, if the stationary point is beyond the area of observed

phenotypes, great care should be taken before any conclusions are drawn about the set of "optimal" characters in relation to the observed character composition of the population. Any deviations from the quadratic fit will become amplified as the distance of the calculated optimum from the population mean increases, and therefore, the calculated optimum may be quite far from the actual optimum in this case. If the calculated optimum is within the range of phenotypic expression, however, the confidence region of the stationary point will need to be considered if, for example, the traditional definitions of stabilizing and disruptive selection are to be used (see also Mitchell-Olds and Shaw [1987]).

Despite being a simple transformation, canonical analysis does reveal a cautionary note on the interpretation of the selection coefficients. If the errors of the quadratic selection gradients are large (especially those of γ_{ii}), qualitatively different surfaces may provide a statistically adequate fit to the allowable range of coefficients, and it will be impossible to ascribe a single surface to the coefficients. Therefore, simply bounding the γ_{ii} away from zero does not necessarily imply that the population is experiencing multivariate convex selection (even if canonical analysis confirms this result); one must also simultaneously consider the range of the γ_{ij}.

A further difficulty in the interpretation of the selection coefficients can be encountered if the phenotypic distribution of traits being measured is not multivariate normal. In this case, there can be covariance between the linear and quadratic terms used in the quadratic regression, which can lead to incorrect estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. The correct estimates can be obtained, however, by using a two-step process in which $\boldsymbol{\beta}$ is estimated by using a purely linear regression, and $\boldsymbol{\gamma}$ is estimated using the full quadratic regression (Lande and Arnold, 1983 p. 1218). Having two estimates of $\boldsymbol{\beta}$ can lead to a dilemma when one wants to visualize the selection surface, however. The full regression gives the best quadratic fit to the surface but yields incorrect estimates of $\boldsymbol{\beta}$. One is then left with the choice of using the coefficients of the full regression or the correct selection coefficients when reconstructing the surface. The set of coefficients chosen depends on whether one wants to visualize the best approximation to the surface or the selection coefficients themselves. We argue for the latter (see *Fitness Surfaces* above).

Finally, it should be noted that the SAS statistical package (SAS Institute, 1985) contains a regression procedure (RSREG) that will perform many of the procedures described above. Because of the current rarity of multivariate data sets of sufficient size to yield reasonable estimates of quadratic selection coefficients, we do not present a detailed example of how to perform a canonical analysis. E. Simms (unpubl.), however, provides an application of the canonical analysis of a selection surface using data on plant resistance to multiple herbivores.

Rationale for Multivariate Studies of Nonlinear Selection

Selection studies often analyze directional selection and ignore nonlinear selection. This practice is particularly unfortunate in cases where lifetime fitness has been measured (e.g., Clutton-Brock, 1988). In such cases, if the population is in evolutionary equilibrium, there is no directional selection on heritable characters ($\boldsymbol{\beta} = \mathbf{0}$), and information about selection pressures will reside instead in the coefficients of quadratic selection. Even in nonequilibrium populations or when only components of fitness have been measured, coefficients of quadratic selection can provide a great deal of information. Quadratic coefficients can identify nonlinear relationships between traits and fitness (e.g., intermediate optima) and can show that selection has acted on functionally coupled traits. For instance, Arnold and Bennett (1988) used correlational selection analysis to suggest that characteristic combinations of body and tail vertebral numbers promoted crawling speed in garter snakes. The optimum number of tail vertebrae varied with the number of body vertebrae within a single population. This type of functional analysis can be further supplemented by canonical analysis of the selection or performance surface. The pattern of multivariate correlational selection is revealed in the transformation matrix, **M**, with traits experiencing correlational selec-

tion with respect to one another loading positively on the same principal axis (Fig. 4). The patterns thus revealed can then be used to generate hypotheses which could be tested on functional or other grounds (see also Arnold [1983]). Finally, quadratic selection gradients are of interest because of their role in the dynamical equations for the evolution of genetic variances and covariance [Eq. (2); Lande, 1980, 1984].

Visualizing Selection

We have pointed out that a series of univariate or even bivariate pictures may give a misleading representation of selection. Given this circumstance, one could settle for a complete set of selection coefficients that describe multivariate selection, or one could go further and attempt to visualize the selection surface implied by the coefficients. There are several reasons for taking this additional step. First, visualization can provide a compact summary of multivariate selection. Five coefficients are needed to describe selection on two traits, nine coefficients are needed for three traits, and 14 coefficients are needed for four traits. In contrast to the difficulty of digesting such series of coefficients, we can represent selection on two traits as a two-dimensional surface (e.g., Fig. 4), selection on three traits as a nested series of three dimensional solids (see below), and selection on four traits by animating the three-trait representation. All of these graphical techniques are now practical on personal computers. Second, visualization may be useful in comparative studies of selection. For example, the geometric similarity of selection surfaces may not be readily apparent from inspection of selection coefficients alone. Third, visualization may also be useful in planning experimental work. Using the selection surface, one can easily predict how manipulations of trait combinations will affect fitness or its components (Box and Draper, 1987).

Given that one wants to visualize selection, there are still some difficulties to overcome. If selection is only being measured on two traits, one could simply use Equation (3) to regenerate the fitted surface (using a contour-plotting program, for example). If more than two traits are involved, however, correlational selection may complicate the interpretation of the surface. Simply drawing surfaces for pairs of traits will be misleading if there is strong correlational selection between the traits being drawn and those that are not. In addition to the problem of simply drawing the surface, any interpretation of the selection gradients could be incorrect if correlational selection is acting. Several approaches to this difficulty become clear when the problem is viewed in light of canonical analysis.

Because the canonical coefficients provide an orthogonal description of the surface, any drawing of a pairwise combination of transformed variables will yield a correct representation of that aspect of the multidimensional surface. However, the canonical space will almost never be the same as the original character space, and surfaces in the original space may be desired (e.g., for functional interpretation). This problem is alleviated somewhat if canonical analysis indicates that one of the canonical axes is essentially parallel to that of a particular trait. This trait can then be safely drawn in its original space with little chance of misrepresentation. In addition, traits may cluster into groups that experience similar patterns of correlational selection. If selection on such groups of traits is independent of selection on other traits (as indicated by the loading patterns from canonical analysis), then these groups can be drawn together in the original space. For a group consisting of three traits, this can be accomplished by drawing multiple two-dimensional sections of the surface with the third trait varying across graphs or by drawing all three traits simultaneously as a three-dimensional contour drawing of a four-dimensional surface (Fig. 5).

As with selection on two traits, the number of possible quadratic fitness surfaces that result from selection on three traits is limited. The possibilities are shown in Figure 5. In each graph the point of maximum fitness is at the origin (except Fig. 5g, h, in which there are no fixed maxima), and the contours of the surface enclose this maximum like the layers of an onion enclose its center (Box and Draper, 1987 p. 330). Surfaces created by four or more traits cannot

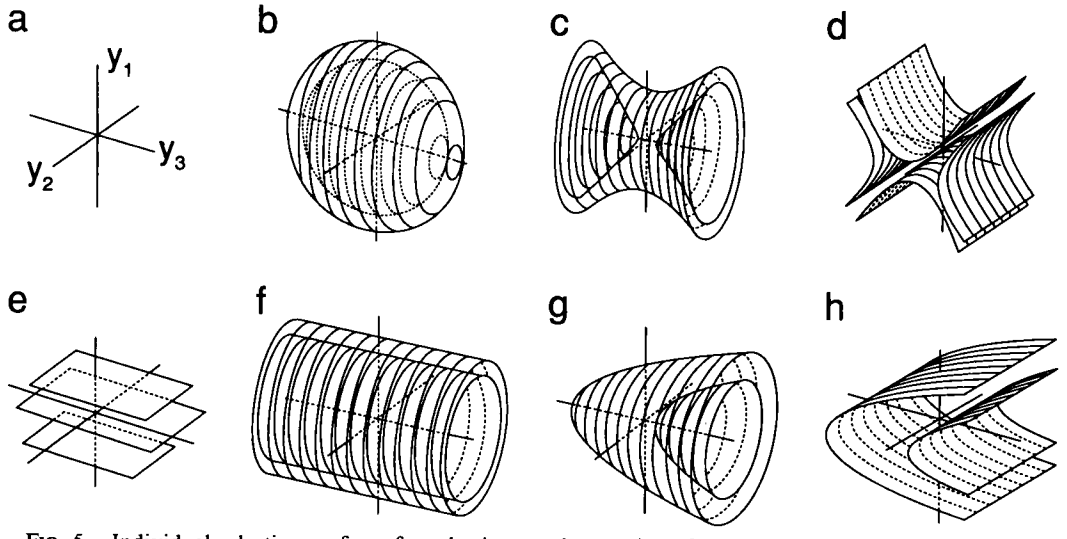

FIG. 5. Individual selection surfaces for selection on three traits (after Box and Draper [1987 fig. 11.3]). Figures are contour representations of a four-dimensional selection surface, fitness being the fourth axis. All points on the surface of a solid have the same fitness. Surfaces are represented in their canonical spaces. The point of maximum fitness is at the origin in b and c. There is a line of maximum fitness in d and f, and a plane of maximum fitness in e. a) Axis labels; b) convex selection on all three traits $(\lambda_1, \lambda_2, \lambda_3) = (-,-,-)$; c) convex selection on traits y_1 and y_2, concave selection on y_3 $(-,-,+)$; d) convex selection on y_1, concave selection on y_3, no selection on y_2 $(-,0,+)$ (compare with Fig. 4b); e) convex selection on y_1, no selection on y_2 or y_3 $(-,0,0)$ (compare with Fig. 4c); f) convex selection on y_1 and y_2, no selection on y_3 $(-,-,0)$ (compare with Fig. 4a); g) convex selection on y_1 and y_2, strong directional selection on y_3 ($\theta_3 > 0$; therefore the center of the system is undefined, i.e., it is a rising ridge) $(-,-,0)$; h) convex selection on y_1, no selection on y_2, strong directional selection on y_3 $(-,0,0)$ (compare with Fig. 4d).

be explicitly visualized, but apart from actually drawing the surfaces, canonical analysis allows one to characterize the entire multidimensional surface by inspection of the canonical coefficients and the transformation matrix. For studies involving many traits, combinations of the above techniques should help to elucidate the form of selection that is operating.

ACKNOWLEDGMENTS

We are grateful to J. H. Willis for discovering the utility of response-surface analysis and for stimulating discussions. M. Kirkpatrick and R. Thompson also pointed us towards canonical analysis. We thank B. Charlesworth, E. Simms, R. Lande, D. Townsend, and M. Wade for helpful discussions. K. Crider, R. Lande, T. Mitchell-Olds, T. Price, M. Wade, G. Wagner, and J. Willis provided numerous helpful comments on the manuscript. This work was supported by NIH grant 1-RO1-GM-35492-01 and NSF grant 85-06766 to S.J.A., and an NSF Graduate Fellowship to P.C.P.

LITERATURE CITED

ARNOLD, S. J. 1983. Morphology, performance and fitness. Amer. Zool. 23:347–361.

ARNOLD, S. J., AND A. F. BENNETT. 1988. Behavioural variation in natural populations. V. Morphological correlates of locomotion in the garter snake *Thamnophis radix*. Biol. J. Linn Soc. 34:175–190.

AYALA, F. J., AND J. W. VALENTINE. 1979. Evolving: The Theory and Processes of Organic Evolution. Benjamin Cummings, Menlo Park, CA.

BERENBAUM, M. R., A. R. ZANGERL, AND J. K. NITAO. 1986. Constraints on chemical coevolution: Wild parsnips and the parsnip webworm. Evolution 40: 1215–1228.

BOX, G. E. P. 1954. The exploration and exploitation of response surfaces: Some general considerations and examples. Biometrics 10:16–60.

BOX, G. E. P., AND N. R. DRAPER. 1987. Empirical Model-Building and Response Surfaces. Wiley, N.Y.

BOX, G. E. P., AND J. S. HUNTER. 1954. A confidence region for the solution of a set of simultaneous equations with an application to experimental design. Biometrika 41:190–199.

BOX, G. E. P., AND K. B. WILSON. 1951. On the experimental attainment of optimum conditions. J. Roy. Stat. Soc. B 13:1–38.

BOX, G. E. P., AND P. V. YOULE. 1955. The exploration and exploitation of response surfaces: An example of the link between the fitted surface and the basic mechanism of the system. Biometrics 11: 287–323.

CAVALLI-SFORZA, L. L., AND M. W. FELDMAN. 1981. Cultural Transmission and Evolution: A Quantitative Approach. Princeton Univ. Press, Princeton, NJ.

CLUTTON-BROCK, T. 1988. Reproductive Success. Univ. Chicago Press, Chicago, IL.

DRAPER, N. R., AND H. SMITH. 1981. Applied Regression Analysis, 2nd Ed. Wiley, N.Y.

ENDLER, J. A. 1986. Natural Selection in the Wild. Princeton Univ. Press, Princeton, NJ.

FALCONER, D. S. 1981. Introduction to Quantitative Genetics, 2nd Ed. Longman, N.Y.

GRANT, V. 1977. Organismic Evolution. Freeman, San Fransisco, CA.

KIMURA, M. 1956. Rules for testing stability of a selective polymorphism. Proc. Nat. Acad. Sci. USA 42:336–340.

———. 1983. The Neutral Theory of Molecular Evolution. Cambridge Univ. Press, Cambridge, U.K.

LANDE, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30: 314–334.

———. 1979. Quantitative genetic analysis of multivariate evolution applied to brain : body size allometry. Evolution 33:402–416.

———. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

———. 1984. The genetic correlation between characters maintained by selection, linkage and inbreeding. Genet. Res. 44:309–320.

———. 1988. Quantitative genetics and evolutionary theory, pp. 71–84. *In* B. S. Weir, E. J. Eisen, M. M. Goodman, and G. Namkoong (eds.), Proceedings of the Second International Conference on Quantitative Genetics. Sinauer, Sunderland, MA.

LANDE, R., AND S. J. ARNOLD. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

LERNER, I. M. 1954. Genetic Homeostasis. Dover, N.Y.

LUSH, J. L. 1945. Animal Breeding Plans. Iowa State Univ. Press, Ames.

MANLY, B. F. J. 1985. The Statistics of Natural Selection. Chapman and Hall, N.Y.

MCGREGOR, P. K., J. R. KREBS, AND C. M. PERRINS. 1981. Song repertoires and lifetime reproductive success in the Great Tit (*Parus major*). Amer. Natur. 118:149–159.

MEAD, R., AND D. J. PIKE. 1975. A review of response surface methodology from a biometric viewpoint. Biometrics 31:803–851.

MITCHELL-OLDS, T., AND R. G. SHAW. 1987. Regression analysis of natural selection: Statistical inference and biological interpretation. Evolution 41: 1149–1161.

PEARSON, K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlations of organs. Phil. Trans. Roy. Soc. Lond. A 187:253–318.

PRICE, T. D., AND P. T. BOAG. 1987. Selection in natural populations of birds, pp. 257–287. *In* F. Cooke and P. A. Buckley (eds.), Avian Genetics. Academic Press, London, U.K.

PRICE, T. D., P. R. GRANT, H. L. GIBBS, AND P. T. BOAG. 1984. Recurrent patterns of natural selection in a population of Darwin's finches. Nature 309:787–789.

ROBERTSON, A. 1956. The effect of selection against extreme deviants based on deviation or on homozygosis. J. Genet. 54:236–248.

SAS INSTITUTE. 1985. SAS User's Guide: Statistics, Version 5 Ed. SAS Inst., Inc., Cary, NC.

SCHLUTER, D. 1988. Estimating the form of natural selection on a quantitative trait. Evolution 42:849–861.

SCHMALHAUSEN, I. I. 1949. Factors of Evolution: The Theory of Stabilizing Selection. Univ. Chicago Press, Chicago, IL.

SIMPSON, G. G. 1953. The Major Features of Evolution. Columbia Univ. Press, N.Y.

SPIESS, E. B. 1977. Genes in Populations. Wiley, N.Y.

STABLEIN, D. M., W. H. CARTER, AND G. L. WAMPLER. 1983. Confidence regions for constrained optima in response-surface experiments. Biometrics 39:759–763.

THOMPSON, R. 1977. Estimation of quantitative genetic parameters, pp. 639–657. *In* E. Pollak, O. Kempthorne, and T. B. Bailey (eds.), *Proceedings of the International Conference on Quantitative Genetics*. Iowa State Univ. Press, Ames.

TURELLI, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theoret. Popul. Biol. 25:138–193.

———. 1985. Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. Genetics 111:165–195.

WADDINGTON, C. H. 1957. The Strategy of the Genes. Allen and Unwin, London, U.K.

WRIGHT, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. 6th Internat. Congr. Genet. 1:356–366.

———. 1969. Evolution and the Genetics of Populations, Vol. 2. The Theory of Gene Frequencies. Univ. Chicago Press, Chicago, IL.

———. 1977. Evolution and the Genetics of Populations, Vol. 3. Experimental Results and Evolutionary Deductions. Univ. Chicago Press, Chicago, IL.

Corresponding Editor: M. R. Lynch