



The adaptive landscape as a conceptual bridge between micro- and macroevolution

Stevan J. Arnold, Michael E. Pfrender & Adam G. Jones

Department of Zoology, 3029 Cordley Hall, Oregon State University, Corvallis, OR 97331, USA (Phone: (541) 737-4362; Fax: (541) 737-0501; E-mail: arnolds@bcc.orst.edu)

Key words: adaptive landscape, macroevolution, microevolution, phenotypic evolution, quantitative genetics, selection surface, selective line of least resistance

Abstract

An adaptive landscape concept outlined by G.G. Simpson constitutes the major conceptual bridge between the fields of micro- and macroevolutionary study. Despite some important theoretical extensions since 1944, this conceptual bridge has been ignored in many empirical studies. In this article, we review the status of theoretical work and emphasize the importance of models for peak movement. Although much theoretical work has been devoted to evolution on stationary, unchanging landscapes, an important new development is a focus on the evolution of the landscape itself. We also sketch an agenda of empirical issues that is inspired by theoretical developments.

Introduction

Is the ‘modern synthesis’ incomplete? At the center of disenchantments with the neo-Darwinian theory of evolution is the connection between micro- and macroevolution. The term microevolution refers to the processes that lead to phenotypic diversification among arrays of conspecific geographic races or closely related species. Macroevolution, on the other hand, covers processes responsible for the divergence among genera or higher taxa. We favor the view that neo-Darwinian theory can account for both micro- and macroevolutionary patterns (Lande, 1980a; Charlesworth, Lande & Slatkin, 1982). Nevertheless, despite our optimism, we recognize that disenchantment is easy to find in the literature of evolutionary biology. The main complaints fall into two broad categories: (1) claims that microevolutionary processes cannot logically be extrapolated to explain macroevolutionary pattern (Stanley, 1979; Eldredge & Cracraft, 1980), and (2) the idea that important pattern-producing processes operate above the level of populations (e.g., species selection; Rensch, 1959; Vrba, 1983). The conceptual chasm between microevolutionary processes (inheritance, selection, drift)

and macroevolutionary patterns appears to some authors to be deep, wide and unbridgeable. Remarkably, a conceptual bridge was outlined more than 50 years ago by Simpson (1944, 1953) but is neglected by many evolutionary biologists today.

Simpson (1944) boldly used an adaptive landscape to synthesize genetical and paleontological approaches to evolution. In Simpson’s conceptualization a two-dimensional space represents the possible combinations of two phenotypic characters (structural variants). Elevation contours on this space represent population fitness (adaptiveness). Using this phenotypic landscape, Simpson illustrated the concepts of phenotypic variation, selection, immediate responses to selection, long-term evolutionary trends, speciation, and adaptive radiation. No visualization before or since 1944 has been so successful in integrating the major issues and themes in phenotypic evolution.

Topographic simplicity and peak movement are notable features of Simpson’s landscapes. He usually portrayed just one or two adaptive peaks. Peak movement is a second important theme. Simpson modeled the tempo and mode of evolution with various patterns of peak bifurcation and movement. In Simpson’s conceptualization the population evolves in relation to a

changing landscape. The model is not one of evolution on a complex but stationary landscape. In recent years, the theoretical literature has explored the issue of peak movement. This change in focus from stationary to evolving landscapes is so profound that it is fair to call it a paradigm shift, yet it has escaped the notice of many evolutionary biologists. We shall return to the themes of topographic simplicity and peak movement later in our discussion.

The landscape under discussion should not be confused with certain other landscape concepts in the literature of evolutionary biology. The landscape that Simpson used, and which we will explore, is a space of phenotypic characters (Lande, 1976a, 1979). Elevation on this space reflects population-level fitness (adaptation). This phenotypic landscape is a direct descendant of Wright's adaptive landscape (Wright, 1931, 1932, 1945), except that his landscape is a space of gene frequencies (Wright, 1932, 1982; Provine, 1986). In Wright's conceptualization, the landscape is complex (due to epistasis in fitness) and largely stationary. Movement on the Wrightian landscape represents evolutionary change in gene frequencies rather than phenotypic evolution *per se*. Other landscapes are still more distant to the one under discussion. Rice (1998) uses a landscape to model the evolution of phenotypic plasticity in which elevation represents a character and the axes reflect underlying factors. Waddington's (1957) epigenetic landscape is a space of abstract variables that is used to describe the modal developmental tendency and major deviations from it.

In evolutionary biology today Simpson's landscape is not routinely used to motivate empirical work, even though its power has been confirmed and extended by theoretical studies. The theoretical developments are relatively recent, tracing back to Lande (1976a, 1979), and are often couched in the language of multivariate calculus and linear algebra. Simpson's landscape lives and flourishes in these theoretical papers, but rarely is illustrated. Consequently, the idea that Simpson's landscape is the major conceptual bridge between the fields of micro- and macroevolution is unappreciated by many evolutionary biologists. The goal of this article is to give an overview of theoretical developments in the field of phenotypic evolution, especially those that can be visualized with Simpson's landscape. Our thesis is that these results and visualizations could and should guide empirical work in a wide variety of disciplines. Our survey also highlights some directions that need theoretical exploration. The bridge is still under construction.

Some recent works are important companions to our discussion. Hansen and Martins (1996), building on the work of Felsenstein (1973, 1985), have pointed out that the fields of systematics, evolutionary genetics, and comparative biology rest on a common set of equations relating evolutionary pattern (trait variance and covariance among taxa) to process (mutation, selection, drift). Those unifying equations are in turn based on models that relate microevolution to macroevolution. In this review we give a landscape visualization of the various process models that are central to Hansen and Martins' (1996) discussion. Schluter (2000) has used Simpson–Lande landscapes to illustrate the concept of adaptive radiation and to survey the growing empirical literature. We will rely on Schluter's (2000) treatment of adaptive radiations, while extending his discussion of landscapes and how they can be used. To provide a connection to the theoretical literature, while keeping the text free of mathematical notation, we will indicate equations by number in parenthesis. The corresponding mathematical expressions and their attributions are given in the Appendix.

Current conceptualizations of the adaptive landscape

Overview

The adaptive landscape for continuously distributed, phenotypic characters is a surface that relates average fitness to average character values. Although only the one character case is usually portrayed in textbooks, the landscape must be visualized in at least two dimensions to appreciate fully the key concepts. The landscape is more than a theoretical construct – crucial features of this surface can be determined empirically.

In this section, and the ones to follow, we will use a landscape concept in which selection favors an intermediate optimum. There are many reasons for this choice for selection, and chief among them is its firm empirical foundation. This form of selection (sometimes called stabilizing or centrifugal selection) was documented in some of the earliest empirical studies of phenotypic selection (Karn & Penrose, 1951) and has been found in many subsequent studies (Endler, 1986; Kingsolver et al., 2001). Stabilizing selection can produce a persistent equilibrium, a result that appeals to many naturalists, in contrast to linear selection regimes under which populations are perpetually subjected to directional change. Long-maintained

stabilizing selection can explain such diverse phenomena as character canalization, geographic variation, and transgressive segregation in second-generation hybrids resulting from a wide cross (i.e., the appearance of variants outside the range of the parental populations) (Mather, 1941, 1943; Schmalhausen, 1949; Wright, 1968; Rieseberg, Archer & Wayne 1999). Furthermore, directional selection can be accommodated in theoretical work by simply shifting the selective optimum away from the character mean (Lande, 1976a). Although we have chosen to illustrate the landscape concept with stabilizing selection, other modes of selection are feasible and the sections that follow could be revisited using those alternative selection modes. In particular, a recent review of the selection literature discovered that instances of disruptive selection were as common as stabilizing selection (Kingsolver et al., 2001), a result that challenges our emphasis on stabilizing selection. Univariate studies of selection predominated in that review, so the jury is still out on whether the adaptive landscape is commonly hill, pit, saddle or ridge-shaped (Phillips & Arnold, 1989).

The adaptive landscape for a single character: the march of the frequency distribution

The adaptive landscape for a single character under stabilizing selection can be represented by a dome-shaped curve. If the mean of the character is situated some distance from the apex (optimum) of the curve, the population experiences directional selection that will tend to shift the mean toward the optimum if the character is heritable. If heritability is constant, the amount of change across generations is proportional to the distance to the optimum (Lande, 1976a). Evolution towards a stationary optimum is rapid at first, decelerates as the mean approaches the optimum, and then ceases entirely when the mean coincides with the optimum (Figure 1), (1). In other words, the frequency distribution marches until it lies under the peak of the landscape. This march corresponds to a progressive increase in average fitness, ceasing when the population achieves a fitness maximum directly under the peak (Lande, 1979), (2). In ecological terms, the movement of the optimum away from the character mean might correspond to a change in climate, resources, predators or change in any other set of conditions that induces directional selection. The movement might happen instantaneously and then cease. An ecologically more plausible circumstance is that the peak movement con-

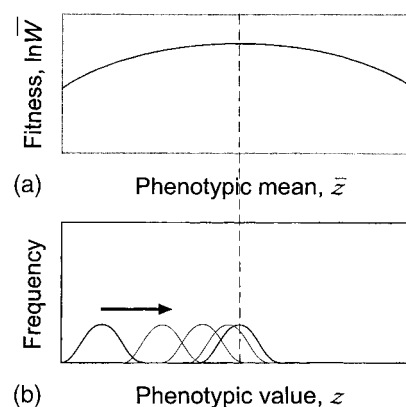


Figure 1. An adaptive landscape for a single character under stabilizing selection (a). The natural logarithm of mean population fitness is shown as a function of phenotypic mean. Evolution of the distribution of phenotypic values in response to the adaptive landscape (b).

tinues over a period of generations so that for some period of time the population chases an ever-moving optimum. Thus, during the Pleistocene, most periods of, say, a hundred generations might have been characterized by progressive change in temperature that in turn induced continued change in a host of ecological variables affecting fitness. For any particular character this progressive change translates into an optimum that moves steadily away from the mean in the same direction. Colonization of a new environment can also create a situation in which the trait mean is some distance from the optimum with resulting rapid evolution. Colonization of new hosts, new spawning habitats and environments with new predators are examples (Via, 1991; Reznick et al., 1997; Feder, 1998; Hendry et al., 2000).

A stationary adaptive landscape for two characters: the simplest case of multivariate evolution

The adaptive landscape for two characters under stabilizing selection can be represented by a hill-shaped topography. The optimum is represented by the crest of the hill. If the bivariate character mean is located some distance from the optimum, the population experiences directional as well as stabilizing selection. The strength of directional selection corresponds to the direction of steepest uphill slope from the character mean on the adaptive landscape (Lande, 1979), (3). This vector can be resolved into two components or selection gradients, corresponding to the two character dimensions (Figure 2). A steeply sloping hill represents strong directional selection, whereas a

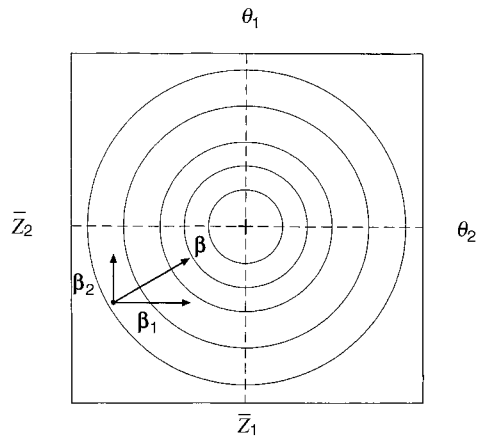


Figure 2. An adaptive landscape for two characters under stabilizing selection with no correlational selection. The vectors show the forces of directional selection that act on a population whose bivariate mean is at the position of the small dot (at the base of the vectors). θ_1 and θ_2 denote the position of the optima for the two characters.

weakly sloping hill corresponds to weak directional selection. The intensity of stabilizing selection corresponds to the curvature of the landscape, evaluated at the character mean (Lande & Arnold, 1983), (4). Landscape curvature has three aspects. The first two aspects are stabilizing selection corresponding to the two character dimensions and can be thought of as curvature of the hill, viewed in slices parallel to the two character axes. Strong stabilizing selection corresponds to strong curvature, weak stabilizing selection corresponds to weak curvature. The third aspect of curvature concerns the orientation of the hill, and is called correlational selection. If the long axis of the hill is parallel to one of the character axes, there is no correlational selection. A hill with an upward-tilting axis is said to impose positive correlational selection (Figure 3). A hill with a downward-tilting axis is said to impose negative correlational selection. These three aspects of curvature can be represented by a matrix, the so-called γ -matrix (5), (Lande & Arnold, 1983; Arnold, 1992).

A general principle of evolution on the adaptive landscape is that the population mean tends to move uphill (Lande, 1979), (2). If the landscape is of a simple type shown in Figure 3, with one adaptive peak, this principle means that the population mean will tend to evolve towards equilibrium on that peak, although genetic drift can cause departures from upward movement and equilibrium. Additional exceptions to this fitness maximization principle will be discussed

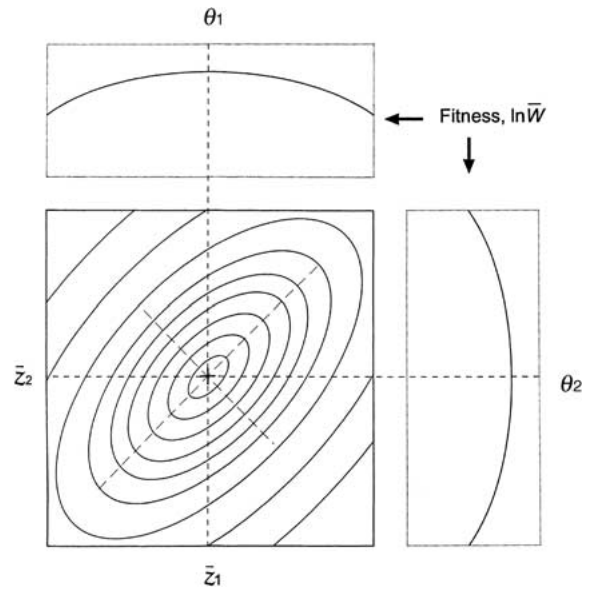


Figure 3. An adaptive landscape for two characters under stabilizing and positive correlational selection. Same conventions as in Figure 2. The long and short axes of the landscape are the principal components (eigenvectors) corresponding to the two largest eigenvalues of the landscape.

in a later section. Although the population tends to evolve uphill under frequency-independent selection, the mean generally does not evolve in the direction of greatest improvement in average fitness, the direction of which is given by the directional selection gradients. Instead the population will take a curved path that deviates from this direction of greatest improvement (Lande, 1979, 1980a). Such curved paths arise from unequal genetic variances (i.e., some characters having greater genetic variance than others) and from genetic covariance. These genetic parameters affect both the rate and direction of evolution and are represented by a variance-covariance matrix known as the \mathbf{G} -matrix.

The crucial genetic parameters that affect rate and direction of evolution can be visualized using a set of axes called principal components (5). A population capable of rapid evolution will have abundant genetic variation in both principal components (5). We can represent this abundance with an ellipse that is broad in both axes (Figure 4(a)). The width of the ellipse in slices through its center, parallel to the two character axes, is proportional to the genetic variances that exist in those characters. The greater the genetic variance for a character, the faster it can evolve. The ellipse also has an orientation with respect to the two character axes. This orientation can

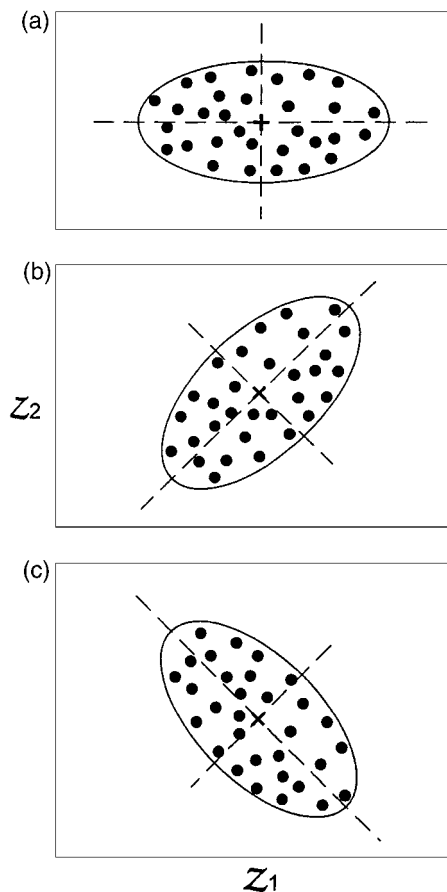


Figure 4. Bivariate distributions of breeding (additive genetic) values representing different patterns of genetic variance and covariance. (a) Large genetic variance in trait 1, small genetic variance in trait 2 and no genetic covariance. (b) Positive genetic covariance. (c) Negative genetic covariance.

be represented by the long axis of the ellipse (i.e., the eigenvector corresponding to the largest eigenvalue of the \mathbf{G} -matrix). When the long axis is parallel to one of the character axes (Figure 4(a)), there is no genetic covariance between the two characters. An upward-tilting axis corresponds to positive genetic covariance (Figure 4(b)); a downward-tilting axis corresponds to negative genetic covariance (Figure 4(c)). The tilt of the genetic variation axis (genetic covariance) can greatly affect the population's response to the adaptive landscape.

One way of appreciating the effect of genetic covariance is to ask, 'under what conditions will the population evolve in the same direction as the direction specified by the directional selection gradients?' In other words, under what conditions will the population evolve in a straight line rather than on a curved path

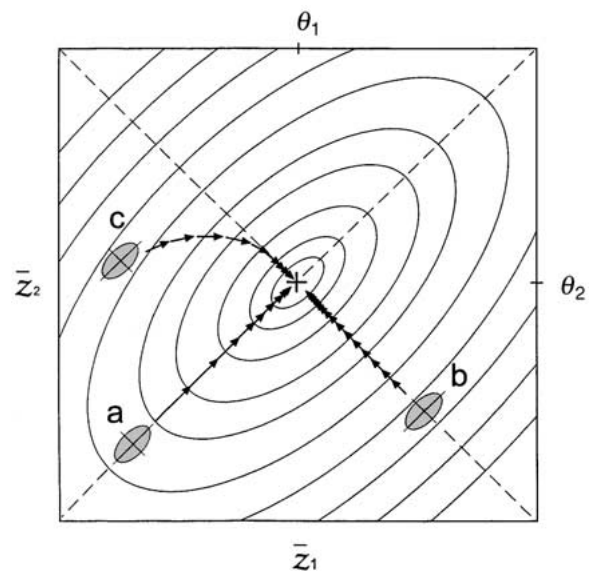


Figure 5. Evolution on an adaptive landscape depends on the alignment of the axes (principal components) of genetic variation (shaded ellipses) with the axes (principal components) of the adaptive landscape. Evolution follows straight trajectories when major (lower left) or minor (lower right) axes are aligned. In general, axes are out of alignment (upper left) and evolution follows a curved trajectory. The small ellipses around each of the three population means represent genetic variation around each mean (the eigenvectors and eigenvalues of the \mathbf{G} -matrix) and hence are on a different scale of measurement.

that deviates from the direction of greatest improvement in fitness? The answer is that the population will evolve in a straight line when an axis of genetic variation is aligned with an axis of the landscape and hence with the selection gradient. Evolution will be rapid when the major axis (first principal component) is aligned (Figure 5(a)) and slow when the minor axis (second principal component) is aligned (Figure 5(b)). The population will tend to evolve on a curved path whenever the axes of genetic variation and the landscape are out of alignment (Figure 5(c)).

The adaptive landscape in more than two character dimensions

No additional concepts are needed to specify an adaptive landscape for three or more characters, although the landscape does become progressively more difficult to visualize. The landscape for three characters under stabilizing selection, for example, can be visualized as a nested series of spheres or ellipsoids (Phillips & Arnold, 1989).

Models for the evolution of the optimum: macroevolution pattern from microevolutionary process

Overview

Models for the movement of the peak of the adaptive landscape can be used to describe changing ecological opportunity or temporal change in the environment. The most successful models of this kind make predictions about macroevolutionary pattern from the microevolutionary processes of selection, drift and inheritance. Because these models characterize expected evolutionary patterns in statistical terms, they can be used to test alternative visions of the adaptive landscape. We will review these characterizations and outline progress in using them to construct tests for the causes of evolutionary pattern.

The following discussions of models follow a common format.

- (1) A microevolutionary process model is specified. The specification corresponds to assumptions about the adaptive landscape and inheritance.
- (2) Using this model, we present expressions for the expected variances and covariances among populations or other taxa for a set of traits. This approach assumes a phylogeny for the populations. The phylogeny that makes the models most tractable is a star, in which all populations diverge simultaneously from one ancestor and are viewed after some number of generations.
- (3) Alternative models for process are distinguished by comparing their predictions about pattern (among-population variances and covariances).

To test alternative process with real data, the assumption of a star phylogeny can be relaxed using an estimate of phylogeny with branch lengths that are proportional to elapsed time. Using those branch lengths, the expected pattern across the whole phylogeny can be calculated (Hansen & Martins, 1996; Martins & Hansen, 1996; Hansen, 1997).

Evolution on static landscapes

The following models consider evolution on landscapes that do not change over time. While less realistic than other models in which the adaptive landscape itself evolves, these models provide important null hypotheses against which to test empirical observations.

Multivariate drift: evolution on a flat landscape

The default topography for the adaptive landscape is a flat and level surface, the drift landscape. Selection does not affect the evolution of the population mean, which instead evolves in a trajectory that can be described by Brownian movement. Because of the simplicity of drift, we can predict the average evolutionary outcome. That outcome depends on elapsed time, effective population size, and the matrix of genetic variances and covariances (the **G**-matrix; Lande, 1976a, 1979). Imagine a set of replicate populations derived instantaneously from the same ancestral population and diverging under drift alone. After any number of generations the expected character mean of all these descendant populations will be the same as the original, ancestral mean. The expectation is that drift will not change the character mean. Even though the average population should have the same mean as its ancestor, divergence among populations in mean can be appreciable and will show a characteristic pattern. The variance–covariance matrix for the means of descendant populations will be proportional to the **G**-matrix (Figure 6). Variance among populations will also be proportional to the number of elapsed generations and inversely proportional to average effective population size (6). Thus, on a drift landscape, the **G**-matrix and effective population size encapsulate, respectively, the microevolutionary processes of inheritance and drift. The equation (6)

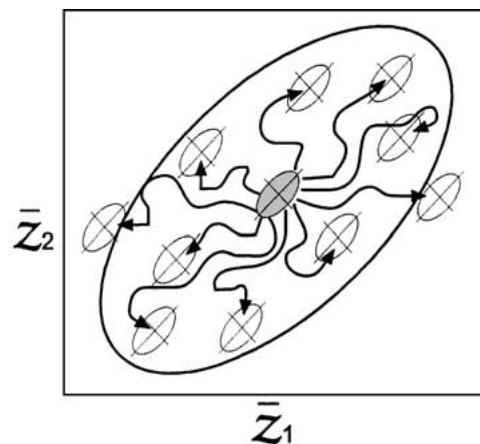


Figure 6. Bivariate drift on a flat adaptive landscape. The small ellipse at the center represents the **G**-matrix of the ancestral population. The large, outer ellipse represents 95% confidence ellipse for the means of replicate, descendant populations. Solid curved lines show representative evolutionary trajectories. Other conventions as in Figure 5.

relating these processes to divergence in means constitutes the bridge between microevolutionary process and macroevolutionary pattern.

The evolutionary trajectory of a small population on a drift landscape is especially chaotic. The small size of the population causes sampling effects that induce random changes in genetic means, variances and covariances. Phillips, Whitlock and Fowler (2001) provide a dramatic example of \mathbf{G} -matrices varying under drift.

Other kinds of evolution on featureless landscapes

The simplest departure from the drift landscape is a surface that is flat but tilted in one or more character dimensions. A population on such a landscape experiences only directional selection. To see the predicted connection between microevolution and macroevolution, imagine the same scenario as before: a set of large descendant populations derived at the same instant from the same ancestral population. All descendants experience the same average tilt to their adaptive landscape, but the tilt in each population fluctuates. Those fluctuations may covary among populations, but they are independent of trait means within populations. If we take a snapshot of these populations after some number of generations (t), the variance-covariance matrix describing their means will be proportional to the product of the \mathbf{G} -matrix, variation in landscape tilt, and the number of elapsed generations (Felsenstein, 1988; Zeng, 1988), (7). If the descendant populations are small in size, genetic drift of the trait mean cannot be ignored. In this case, the variance among population means after t generations has two sources: one due to fluctuation in selection, the other due to sampling errors arising from finite population size (Hansen & Martins, 1996), (8). As in the case of drift, these simple models of pure directional selection yield equations that bridge the chasm between microevolution and macroevolution.

The goal of these simplified models is not to provide a description of how macroevolutionary patterns are actually generated in the real world. The significance of the drift and directional selection models is that they provide benchmarks against which we can test for the signature of selection in nature. If we can refute the predictions of the drift model, we have good evidence that some form of selection is producing pattern. Thus, Lynch (1990) was able to reject a drift model by showing that there was too little variance among species in morphological traits.

If, at the next step, we can reject a model of pure directional selection, we have evidence that more complicated forms of selection are responsible for evolutionary pattern. Equations derived from the models not only bridge between micro- and macroevolution, they provide tests for progressively more complicated kinds of selection. Different kinds of selection are recognized by their characteristic, pattern-producing features. It is important to realize that the tests just described depend on the assumption that the \mathbf{G} -matrix is homogeneous across populations and species, an assumption that we will examine in a later section.

Evolution on a curved but unchanging landscape

An attractive feature of curved landscapes is that they allow the possibility of local equilibria. In other words, each population can remain under selection and yet evolve towards a stable outcome, unless the adaptive peak is constantly in motion. Because those stable outcomes can vary in space and time, they map onto observations of geographic variation (e.g., ecotypic variation).

Lande (1976a, 1979) described a simple model of this kind in which the adaptive landscape consists of a single adaptive peak, Gaussian in shape. The location of the peak and the curvature of the landscape are constant through time. Lande's results are general in the sense that the effects of finite population size are included, but they are approximate because the \mathbf{G} -matrix is assumed to be constant. The mean of a large population tends to evolve towards the peak and will eventually reach it, unless conditions change. Drift arising from finite population size causes the mean to deviate from the optimum. Departure from the optimum generates directional selection towards the optimum; the greater the departure, the stronger the selection. Stabilizing selection generates a restraining force that can be likened to a rubber band, pulling the mean towards the optimum (Hansen & Martins, 1996). For a single character, this restraining force is proportional to the product of genetic variance and the curvature of the landscape (9). The more genetic variance and the greater the curvature of the peak, the more the mean is pulled towards the optimum. The prediction of this model is that the variation among descendant populations will be much less than under pure drift. Because among population variation decays as an exponential function of time and the restraining force, with even weak stabilizing selection there may be no appreciable variation.

Adaptive landscapes with two stationary peaks (Felsenstein, 1979) have been used to model speciation and punctuated equilibria. These models are the phenotypic analog of Wright's shifting balance theory (Wright, 1932, 1940) in the sense that a population can become trapped on a peak that is lower than an adjacent peak. The models are commonly called 'peak shift' models, but the peaks do not move. Instead, the population mean shifts from one peak to the other. Peak shift models incorporate random genetic drift as a mechanism that enables the population to escape from a local peak. Lande's (1986) review summarizes results from several models, some of which are surprising. Under a wide range of conditions, populations show a pattern of relative stasis in which the phenotypic mean erratically drifts in the immediate vicinity of one peak for a long period of time. The average length of this period of relative stasis (the expected time until the population shifts to the other peak) is long if the population is large, the original peak is high, and the valley between the peaks is deep (11). Surprisingly, the expected time until a shift is almost independent of the distance between the two peaks. If we focus on those rare events in which the mean shifts to the second peak, we find that the transit down to the valley – against the force of directional selection – is just as fast as the transit back up to an adaptive peak. This unexpected result is a consequence of sampling rare events in which peak shifts occur. Only especially rapid instances of drift are included in the sample, and in these the speed of downhill transit is just as rapid as the episode of uphill transit. Despite their simplicity, peak shift models produce an evolutionary tempo in which long periods of relative stasis are punctuated by rapid evolutionary transitions. Lande (1986) and other authors of peak shift models have argued persuasively that the conditions underlying these models are more plausible than the set of assumptions invoked by Gould and Eldredge (1977). Models with two stationary peaks have also been used to explore the interaction between selection and gene flow. In such models, gene flow from an adjacent population can cause a population to equilibrate downslope from its adaptive peak (Garcia-Ramos & Kirkpatrick, 1997; Hendry, Day & Taylor, 2001).

Evolution of curved landscapes

In all the preceding models, we imagined that the adaptive landscape was invariant through time. We now consider the possibility that the landscape itself can

change. One important kind of change is caused by density-dependent selection, which can cause a peak to flatten as the population evolves towards it (Brown & Vincent, 1992; Schluter, 2000). In this section, however, we will focus on change that involves the position of the optimum. A simple kind of landscape evolution is for the position of the optimum to change while the curvature and orientation of the surface remains constant. We will discuss five models for peak movement that all share the characteristic that peak shape and orientation remain constant.

Random movement

The position of the optimum might change stochastically. One tractable model assumes that the optimum shows erratic movement (Figure 7(a)), modeled as a stationary time series (Slatkin & Lande, 1976; Bull, 1987; Felsenstein, 1988; Charlesworth, 1993a, b; Lynch & Lande, 1993; Hansen & Martins, 1996; Lande & Shannon, 1996). The process is stationary in the sense that the parameters describing the time series do not change through time. The peak fluctuations may be independent from one generation to the next or autocorrelated. In such models the evolutionary response of the mean to the moving optimum is an increasing function of the magnitude of the movement, the amount of genetic variance in the direction of the shift and the curvature of the adaptive landscape (12). A set of replicate descendant populations will show trait variation that is proportional to the variance in peak movement, but that variation will decay through time as a negative exponential function (Hansen & Martins, 1996), (13). A challenge in such models is to relate the matrix that controls random movement of the peak to ecological processes. Although the connection is clear in particular cases, a general relationship seems elusive. For example, random peak movement might correspond to erratic fluctuation in climate and other ecological variables.

Selectively constrained movement

A special case of the model just considered is the situation in which the movement of the optimum is constrained by the same forces that influence selection within populations. This prospect seems especially likely for suites of characters governed by biomechanical laws. A model that captures this feature is one in which peak movement is specified by a matrix that resembles the matrix characterizing selection

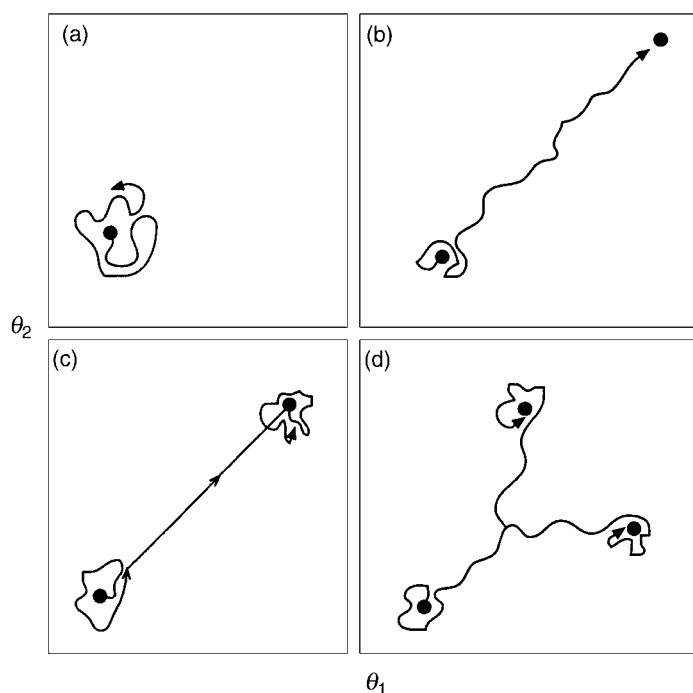


Figure 7. Models for the movement of the optimum of an adaptive landscape for two characters under stabilizing selection. (a) Random, stochastic movement of the optimum. (b) Constant rate and direction of movement of the optimum, after a period of relative stasis. (c) Episodic movement of the optimum separated by periods of relative stasis. (d) Divergence in optima, corresponding to ecological speciation.

acting within populations (14). The situation is analogous to the drift landscape in which a matrix that is proportional to the \mathbf{G} -matrix governs the drift of the mean. One possibility is that a matrix that is inversely proportional to the $\boldsymbol{\gamma}$ -matrix (which specifies the curvature and orientation of the adaptive landscape) governs movement of the optimum. Under this model the optimum undergoes random movement, but that movement is selectively constrained. A prediction of this model is that the pattern of character means will be aligned with the axes of the adaptive landscape (Figure 8). If the $\boldsymbol{\gamma}$ -matrix is the same across populations, the variance–covariance matrix for the means of replicate, descendant populations will be proportional to the inverse of the $\boldsymbol{\gamma}$ -matrix. That among-population variation will, however, decay as an exponential function of time. The weaker the stabilizing selection, the more variation will be retained at any given time. Conversely, strong stabilizing selection will hasten the loss of among-population variation (14).

Constant rate and direction of movement

The optimum might move in characteristic direction at a constant rate (Figure 7(b))(Charlesworth, 1993a,b;

Lynch & Lande, 1993; Lande & Shannon, 1996). This kind of deterministic model might correspond to a constant change in climate, which translates into a steady change in selection pressure. Models of this kind can account for long-sustained evolutionary trends that happen in parallel in multiple lineages. Kurtén (1959), for example, discusses such a pattern in the evolution of mammalian body size during the Pleistocene. On a shorter timescale, this model predicts rapid, evolutionary response of the kind that has been documented in response to anthropogenic changes in the environment (Thompson, 1998; Hendry & Kinnison, 1999; Reznick & Ghalambor, 2001). Depending on the rapidity of change in selection pressures, these rapid responses might also correspond to the following model.

Episodic movement

The optimum might remain relatively constant and then rapidly move to a new position (Figure 7(c)). In Gould and Eldredge's (1977) terminology, peak stasis might be punctuated by periods of rapid movement. Unlike Gould and Eldredge's model of punctuated equilibrium, however, we are not supposing that

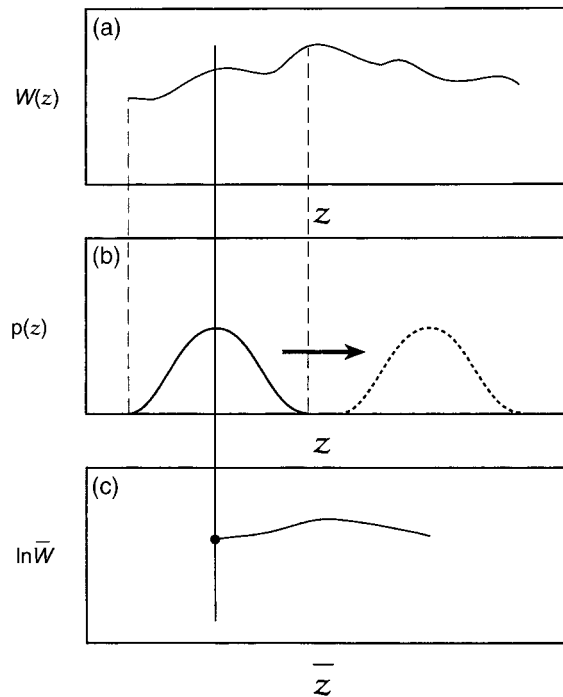


Figure 8. The relationship between the individual selection surface (a) and the adaptive landscape (c). The relationship between these two surfaces can be visualized by averaging the individual selection surface (a) over the phenotypic distribution (solid curve in middle panel) in the interval between the two vertical dashed lines. That averaging produces the solid dot shown in the lower panel. Sliding the phenotypic distribution to the right, and repeating the averaging process at each new position, produces the curve shown in the lower panel, the adaptive landscape.

speciation accompanies periods of peak movement. Our model addresses change within lineages rather than cladogenesis. Episodic movement of the peak might correspond to sudden invasion by a competing species or predator, geological or climatic cataclysm (volcanic eruption, meteor impact, etc.), colonization of a new environment, or anthropogenic change in the environment.

Hansen and Martins (1996) have modeled changes in phenotypic mean that can be related to this kind of within-lineage, episodic peak movement. The parameters in their model, however, are not transparent functions of underlying processes of selection, drift and inheritance. Nevertheless, their model predicts that variation among the means of descendant populations is unlikely to be normally distributed. That result may provide a method that could be used to discriminate between continuous and episodic change in the adaptive landscape. Hansen (1997) models the variation that is expected among species when a slightly

varying optimum moves to a new position. He also provides a worked example showing how phylogeny can be incorporated into the estimation of the two peak positions.

Population extinction is a possible response to movement of the adaptive peak. One way to assess the risk of extinction is to calculate the total genetic load on a population, which is the expected loss in average fitness due to genetic and other factors. Under weak stabilizing selection, the impact of peak movement on genetic load increases as the square of the distance that the peak moves (Lande & Shannon, 1996), (15). Thus, large peak movements are especially likely to cause population extinction. Likewise, rapid movement of the peak may have a profound effect on population persistence. Populations are limited in their rate of evolutionary response to changing conditions by the patterns of genetic variation and covariation for characters under selection. Rapid peak movements may exceed the maximum rate of possible evolution and lead to population extinction (Lynch & Lande, 1993). Population and species differences in genetic architecture will lead to varying capacities to respond to changing landscapes. This process may contribute to macroevolutionary patterns of differential success that some authors have labeled species level selection (Vrba, 1983).

From models for the evolution of the optimum to tests for the causes of evolutionary pattern

A current empirical challenge is to use the models just described to test for alternative causes of evolutionary pattern. From this standpoint, the most successful models are those that relate trait variance/covariance among related taxa to measurable, microevolutionary processes such as selection or inheritance. Thus, the drift model predicts that among species covariance will be proportional to the \mathbf{G} -matrix, whereas the model of selectively constrained peak movement predicts proportionality to the inverse of the \mathbf{Y} -matrix.

Although we have some useful models, much theoretical work remains to be done. A major challenge on the theoretical side is to produce models that yield the most common kinds of patterns in species means. One such common pattern is a correlation in the species means for two variables (e.g., brain weight and body weight). A model of bivariate drift can produce such a pattern (6), but the predicted correlation under that model is likely to be smaller than the one that is commonly observed (Lande, 1979). Models of

flat, tilted landscapes predict patterns of covariance that continually increase through time (7, 8). Models of curved landscapes in which the peak moves randomly about a fixed point predict a steady decay in the covariance of species means (10, 13, 14). None of these models predict a stable pattern of interspecific covariation. In other words, we have an infrastructure for a bridge from microevolutionary process to macroevolutionary pattern, but the construction of the bridge is far from complete. One promising direction might be to build models of peak movement that include ridges, and other channels of movement, that are capable of yielding correlations in species means. Gavrillets (1997) makes a similar point in discussing Wrightian landscapes for genotypic space.

Another way to test the models is to use their predictions concerning the decay in covariance between species as a function of evolutionary distance. Some models predict linear decay, while others predict exponential decay (Hansen & Martins, 1996). Additional theoretical development may facilitate tests of both kinds. For example, if the models can be arranged in a hierarchy so that each successive model differs from a simpler one by a single parameter, then it should be possible to use likelihood ratios to test and reject models in sequence. That goal has not been achieved but it is not far off (Hansen & Martins, 1996).

Empirical characterization of the adaptive landscape

Overview

Key features of the adaptive landscape can be estimated by analyzing variation within populations. In the next sections we describe why such analyses are best pursued as a multivariate problem. We review the connection between a surface that can be estimated from within-population data and the adaptive landscape. Using this connection, we stress the importance of estimating both the curvature and the slope of the adaptive landscape. The parameters that describe slope and curvature are also measures of selection intensity in equations for evolutionary change.

The one character case

The analysis of selection on a single character is deceptively simple. Change in the mean can be used as an indication of directional selection, and change in the variance can be used as an indication of stabilizing

or disruptive selection. The problem with such analysis is that the observed shift in the mean may be a consequence of selection on the character in question (direct selection) or it may be a consequence of selection on correlated characters (indirect selection), (16). Likewise, the observed change in variance (and covariance) may be due to direct or indirect selection (17). Even directional selection on the character in question can cause its variance to contract. For all these reasons, the measurement of selection is a multivariate problem (Lande & Arnold, 1983).

Multivariate selection

The best data for sorting out direct and indirect effects of selection are longitudinal data in which we know the values for a set of phenotypic traits for each individual in a large sample and each individual's fitness. With such data we can use multivariate statistical methods to characterize the surface that relates individual fitness to individual phenotypic values (Lande & Arnold, 1983). This individual selection surface is not the same as the adaptive landscape, but it is closely related to it (Kirkpatrick, 1982; Phillips & Arnold, 1989; Whitlock, 1995; Schluter, 2000, pp. 85–88). Under certain assumptions we can use our characterization of the individual selection surface to estimate key features of the adaptive landscape. In particular, we can use this correspondence between surfaces to estimate the slope and curvature of the adaptive landscape in the vicinity of the population's phenotypic mean.

What is the individual selection surface and how is it related to the adaptive landscape? The individual selection surface is a surface of expected fitness for an individual as a function of the values of its phenotypic characters (18). The relationship between this selection surface and the adaptive landscape is simple if the characters follow a multivariate normal distribution. Under this assumption, the slope of the adaptive landscape is equal to the average slope of the individual selection surface, weighted by the trait distribution (19). The same kind of equivalency holds for the curvatures of the two surfaces (Lande & Arnold, 1983), (20). Thus, if the individual selection surface is wavy, so that slope and curvature vary with position in trait space, the adaptive landscape will have the same average slope and curvature (weighted by the phenotypic distribution), but will be smoother (Figure 8). These equivalencies can be used to estimate the descriptive parameters of the adaptive landscape from

data on individual fitness and trait values. The first step is to characterize the individual selection surface.

Multiple regression can be used to characterize the individual selection surface (Lande & Arnold, 1983; Phillips & Arnold, 1989; Brodie, Moore & Janzen, 1995; Janzen & Stern, 1998). In such an analysis a model is fitted so that individual fitness is predicted from the values of the various characters that are measured. Although such analyses are common in the literature, investigators usually fit only a linear regression. The coefficients that are estimated by linear regression are the average slopes of the surface (β_1 , β_2 , etc.), which are equivalent to the slope of the adaptive landscape if the traits are multivariate normal. To fit a curvilinear regression – so that the curvature and orientation of the surface can be estimated, as well as its slope – one needs to estimate the coefficients for squared and product variables (e.g., z_1^2 , z_1z_2). The coefficients for these curvilinear terms (e.g., γ_{11} , γ_{12}) are the elements in the $\boldsymbol{\gamma}$ -matrix (4). Such a curvilinear regression is known as a quadratic surface (18). Examples of such quadratic surfaces are given in Arnold (1988) and Brodie (1992).

An unfortunate trend in the empirical literature has been to estimate β coefficients and ignore γ coefficients. The trend is unfortunate because $\boldsymbol{\gamma}$ plays an even larger role in evolutionary theory than does $\boldsymbol{\beta}$. The large role of $\boldsymbol{\gamma}$ can be appreciated by scanning the Appendix. Because $\boldsymbol{\gamma}$ describes the curvature of the adaptive landscape, it occurs in many equations that describe the pattern of dispersion of species means, and how that pattern changes through time.

In addition to being descriptors of the adaptive landscape, the β and γ coefficients have another significance. The parameters corresponding to these coefficients are the measures of selection intensity that appear in equations for the evolutionary change in the phenotypic mean and the \mathbf{G} -matrix (Lande & Arnold, 1983; Arnold, 1992). Thus, the quadratic regression just described is a way to estimate parameters of selection that play key roles in evolutionary theory.

If the individual selection surface is the primary object of interest, rather than the adaptive landscape, other methods can be used to describe its features. A limitation of the quadratic regression approach is that it may not accurately represent the individual selection surface, especially if the surface is highly irregular. Projection pursuit regression, a variety of polynomial regression, can be used in such cases (Schluter, 1988; Schluter & Nychka, 1994). A further advantage of these methods is that they do not rest on an assump-

tion of normal trait distributions. Although accuracy of representation and escape from normality are gained with projection pursuit regression, a price is paid. The method does not provide estimates of the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. Quadratic regression can be used to estimate $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, even when the individual selection surface is highly irregular. Thus, when the adaptive landscape, as well as the individual selection surface, is of interest, quadratic regression and projection pursuit regression should be viewed as complementary forms of data analysis.

The analyses just described give a picture of the adaptive landscape only in the immediate vicinity of the mean phenotype in the population. Close to the trait mean we can estimate the slope and curvature of the adaptive landscape. Further away from the mean (e.g., more than a phenotypic standard deviation away), we are much less certain about the shape of the landscape. Three other methods, especially experimental manipulation, can help ameliorate this limitation.

Three other approaches yield information about the adaptive landscape and its history: transplant experiments, experimental manipulation of phenotypes, and retrospective selection analysis. In the transplant approach, a sample of phenotypes from two or more environments is grown in the foreign as well as the native environments, and then fitness is assessed in all individuals (Schluter, 2000). In the most revealing experiments of this kind, crosses are made between all pairs of populations so that first and second generation hybrids can also be grown in all the environments (Rundle & Whitlock, 2001). Such experiments can determine: (a) whether ecological or genetic mechanisms are responsible for any fitness reduction that might occur in hybrids, (b) whether peaks differ in absolute height, and (c) whether populations show highest fitness in their native environments. The latter result is consistent with both a rugged landscape (identical for all populations, but with populations occupying different peaks) and a simple landscape with a history of peak movements.

Experimental manipulation of phenotypes is usually used to test the hypothesis of whether selection might act on a trait, but it can also be used to resolve landscape features. This approach consists of ablating, amplifying or otherwise modifying phenotypes and then assessing fitness in both experimental and control (unmodified) classes of phenotypes. Because only particular traits are altered, leaving a complete background of traits unmodified, this approach can give compelling evidence of selection (Sinervo et al.,

1992; Svensson & Sinervo, 2000). If combinations of traits were modified in a factorial design, this approach (known as response surface analysis in the statistical literature) could help resolve the shape of the individual selection surface. The strength of the approach is that statistical power can be gained at some distance from the phenotypic mean (by increasing the sample size of rare phenotypes). One danger in the approach is that experimental traits can be so exaggerated that interactions with unmodified traits may lead to misleading or even pathological values for fitness.

Retrospective analysis of directional selection can be informative in situations in which peak movement seems likely. Such an analysis requires an estimate of the difference in multivariate means between two sister taxa, at least one estimate of the \mathbf{G} -matrix for the characters in question and the assumption that the \mathbf{G} -matrix has been constant during the period of divergence (Lande, 1979; Turelli, 1988), (21). The net selection gradient estimated from such data measures the minimum amount of directional selection on each character that is required to account for the observed differentiation, given a particular estimate of \mathbf{G} . For example, if one assumes that the means of sister populations are at equilibrium with stabilizing selection and that the difference in means corresponds to a difference in optima, then the net selection gradient estimated for that pair of populations represents the pattern of directional selection that was experienced during the divergence of the optima. For examples of retrospective selection analysis see Price, Grant and Boag (1984), Schluter (1984), Arnold (1988), Dudley (1996), and Reznick et al. (1997).

Open empirical issues

Overview

The landscape world-view highlights many important unresolved empirical issues. Some of these issues have to do with assumptions about the invariance of key evolutionary parameters. We will refer to these as homogeneity issues. Another set of unresolved issues deals with peak movement and hence with ecological connections. We characterize these as alignment issues. Lastly, a set of equilibration issues are concerned with whether the adaptive landscape has an adaptive peak and how closely that peak is approached by the trait mean.

Homogeneity issues

Homogeneity of genetic variances and covariances across related populations and taxa is a convenient simplifying assumption that can greatly facilitate theoretical work and data analysis. Lande (1976b, 1980b) argued that the \mathbf{G} -matrix might equilibrate under the opposing forces of mutation-recombination and stabilizing selection. Just because this assumption has some theoretical justification and is convenient does not mean that it is correct. A number of investigators have adopted this hard-nosed, empirical attitude in comparative studies of \mathbf{G} -matrices (Pfrender, 1998; Arnold & Phillips, 1999; Phillips & Arnold, 1999; Roff, 2000). One trend emerging in these studies is that closely related populations often have very similar, if not identical \mathbf{G} -matrices. Another trend is that the principal axes of the \mathbf{G} -matrix are sometimes conserved even when the matrices are demonstrably not identical or proportional. Current directions in empirical studies of \mathbf{G} -matrices are to make multiple comparisons in a phylogenetic context and to model evolutionary change in those matrices. Another issue is whether homogeneity holds for some kinds of characters more than for other kinds. Thus, life-history characters seem the least likely to maintain homogeneous \mathbf{G} -matrices. These traits experience strong selection that can fluctuate with nearly any kind of ecological change. The landscape for major fitness components, that is, (stage or age specific) viability and fecundity, is almost purely directional with little or no curvature to generate a stabilizing influence. In contrast, traits under stabilizing selection are good candidates for \mathbf{G} -matrix homogeneity.

Homogeneity of phenotypic variances and covariances is also an important, largely unresolved issue. The phenotypic variances and covariances for a set of characters can be assembled into a so-called \mathbf{P} -matrix and that entire matrix can be subjected to statistical analysis and tests. Homogeneity of \mathbf{P} -matrices is further removed from central issues than homogeneity of the \mathbf{G} -matrix, and for a number of reasons \mathbf{P} -matrix structure may not be reflective of \mathbf{G} -matrix structure (Willis, Coyne & Kirkpatrick, 1991), but it is still an important issue. A finding of homogeneous \mathbf{P} -matrices suggests that \mathbf{G} -matrices may be homogeneous and may indicate that the adaptive landscape has long maintained the same curvature (Arnold, 1992). The inferences are indirect, but this disadvantage can be offset by the fact that estimates can be obtained for more populations and taxa than in a study

of \mathbf{G} -matrices (Steppan, 1997a,b; Badyaev & Hill, 2000).

Homogeneity of the adaptive landscape among related populations and taxa is a crucial but largely unexplored issue. Our discussion in this article, for example, has been vastly simplified by assuming that the adaptive landscape commonly shifts its peak position while retaining a characteristic curvature and orientation. Is this assumption valid? Comparative studies that test the proposition of landscape homogeneity are difficult to conduct because extensive data are required for each population to estimate individual selection surfaces. Despite this difficulty, the statistical tools for landscape comparison are already in place. Perhaps the most powerful framework for such comparisons is Flury's (1988) hierarchy of tests for comparing the eigenvalues and eigenvectors of variance-covariance matrices. Flury's approach could be applied to $\boldsymbol{\gamma}$ -matrices. A second challenge is to conduct the comparisons and tests in a phylogenetic framework. So far this goal is also elusive. Nevertheless, the reconstruction of the evolution of $\boldsymbol{\gamma}$ on a phylogeny might be the best way to test the central assumptions of models for peak movement.

Alignment issues

Schluter (1996) has proposed that evolution might often occur along genetic lines of least resistance. The latter phrase refers to the principal axis of the \mathbf{G} -matrix, the direction in character space for which there is the most additive genetic variance or g_{\max} . The basis for Schluter's argument can be seen in Figure 5. When a population approaches a stationary adaptive peak, its evolutionary trajectory will often, but not always (note the trajectory in the lower right), be aligned with g_{\max} . Schluter (1996, 2000) describes tests for alignment of the direction of evolution with g_{\max} and applies them to several case studies.

Evolution might also occur along selective lines of least resistance. Consider the model of selectively constrained peak movement in which the pattern of peak shifts mirrors the pattern of selection within populations (14). In such a model the population trajectory, as the population chases its moving peak, will tend to be aligned with the principal axes of the adaptive landscape. Evolution will tend to occur along selective lines of least resistance. To visualize the model, imagine an adaptive landscape that is Gaussian in all dimensions. We can describe the width of this Gaussian hill with a parameter called ω , which is analogous

to the variance of a bell curve (14). A large ω means that the hill is wide and flat, a small ω means that the hill is narrow and sharply curved in a particular trait dimension. Thus, trait dimensions with the largest ω (smallest γ) correspond to directions of selective least resistance; the peak is most prone to move in those directions. To find the line of selective least resistance, we need to determine the principal components of the $\boldsymbol{\omega}$ -matrix (or the negative inverse of the $\boldsymbol{\gamma}$ -matrix). The largest principal component, corresponding to the direction with the greatest width of the hill, may be called $\boldsymbol{\omega}_{\max}$, and represents the line of selective least resistance. The appropriate test of this hypothesis would be to estimate $\boldsymbol{\omega}_{\max}$ (preferably from multiple populations) and compare that direction with a sample of evolutionary trajectories. Phillips and Arnold (1989) describe how to estimate the principal components of a selection surface.

It may be difficult to distinguish between evolution along genetic lines of least resistance and evolution along selective lines of least resistance. The discrimination is made difficult because a long-term, stable pattern of stabilizing selection will tend to bring the \mathbf{G} -matrix into alignment with the adaptive landscape (Lande, 1980c; Cheverud, 1984; Arnold, 1992). Thus, a logical first step in analysis would be to test for correspondence between g_{\max} and $\boldsymbol{\omega}_{\max}$. If these two directions coincide then the two hypotheses regarding alignment with evolutionary trajectories cannot be distinguished. If g_{\max} and $\boldsymbol{\omega}_{\max}$ are appreciably different, then it might be possible to distinguish between the two hypotheses.

Equilibration issues

If population means are close to their adaptive peaks, then the dispersion of means in character space could be construed as the multivariate pattern of optima. Furthermore, the evolution of the multivariate mean could be equated with peak movement. These equivalencies are most likely to be true for characters with abundant genetic variance on landscapes with strong curvature about a single peak (strong stabilizing selection), a combination of attributes that strongly pulls the mean phenotype towards the peak (Hansen & Martins, 1996). Hansen (1997) develops an approach that makes a weaker assumption about equilibration. In Hansen's model the actual optima of related species deviate from a primary optimum that is an unchanging characteristic of the clade as a whole. Deviations of actual optima from the primary optimum are caused by

small background perturbations in inheritance and selection, as well as by the major ecological features that determine the primary optimum. In this view, much interspecific variation could arise from background factors. One need not assume that all interspecific variation represents variation in peak position.

Using individual selection surfaces, it is possible to estimate the position of the optimum, if it is relatively close to the character mean (Phillips & Arnold, 1989) and so test the hypothesis of equilibrium. Unfortunately, such tests for the location of the optimum are seldom conducted.

The shape of the landscape is an issue that also bears on the assumption of equilibration. Although a single adaptive peak has been assumed in this article, and is often revealed in empirical studies, the adaptive landscape could take many other forms. The possibility of an adaptive ridge should be seriously entertained. Such ridges could produce stable patterns of trait covariance (Emerson & Arnold, 1989; Schluter, 2000). On such an adaptive landscape, selection tends to drive the trait mean phenotype towards the ridge, but the population can move along a level ridge by drift.

Conceptual aspects of the adaptive landscape

Local versus global visions of the adaptive landscape

The adaptive landscape can be viewed from either a local or a global perspective. Although the global view dominates popular discussions, the local view is more in line with both theoretical and empirical developments. By 'local' we mean that the landscape for a particular species is viewed only in the immediate vicinity of its phenotypic mean. A global view takes in a larger expanse of the adaptive landscape that includes multiple species, perhaps an entire radiation. Even though it may be useful in particular situations, the global view is seductive and fraught with dangers. Global views often depict a landscape with multiple adaptive peaks, and sometimes a population is shown as it negotiates this complex topography (e.g., Dawkins, 1996, Figure 5.30; Schluter, 2000, Figure 4.2). This perspective is seductive because it purports to show long range as well as short-term possibilities for adaptive evolution. The global view may be accurate when it describes landscapes that reflect concrete environmental factors, such as the distributions of resources. Thus, in situations in which the

landscape reflects, say, seed size and hardness – and hence the individual selection surface of phenotypes exploiting those seeds – it may portray virtually the entire phenotypic space available to an island community of finches and so may be useful for forecasting evolutionary possibilities.

Even here we should not forget the distinction between the individual selection surface and the adaptive landscape. Because of the smoothing effect of the trait distribution, the adaptive landscape can be smooth, even if the individual selection function is rugged and has multiple peaks (Figure 8). For many kinds of phenotypic characters, however, the landscape beyond the limits of current variation in the population is purely imaginary. For such characters it may be gratuitous to assume that distant peaks exist. The global view will often – perhaps always – be plagued with other serious limitations. The landscape for particular species is bound to change through time, for example. If we use a single landscape to predict the dynamics of multiple species, we cannot account for the possibility that different species may simultaneously experience different changes in elevation at the same point in trait space. By forcing all populations and species to experience the same fitness at equivalent points in trait-space, the global view can seriously distort reality. In contrast, the local perspective does not assume that all species reside on the same landscape. This perspective is also more in line with mathematical characterizations of the adaptive landscape, which are typically restricted to features near a particular phenotypic mean. We can visualize multiple species, while retaining an accurate view of each, by superimposing their adaptive landscapes (e.g., Schluter, 2000, Figure 5.5). The distinction between the two perspectives can be clearly seen in the case of an adaptive radiation in which descendant species occupy different adaptive peaks. In the global view the landscape is necessarily complex with multiple peaks (Figure 9(c)). In the local view each species might have continuously experienced a very simple local topography as it followed a moving peak. At a particular point in time, the superimposed landscapes for the different species would depict the sites of the different peaks in trait space (Figure 9(b)). In other words, the fact that a clade shows diverse phenotypes (Figure 9a) does not force us to adopt the global perspective.

Trait interactions provide another reason to view the landscape as a local phenomenon rather than as an ecological reality that exists separately from the organism and its population. By trait interactions we mean

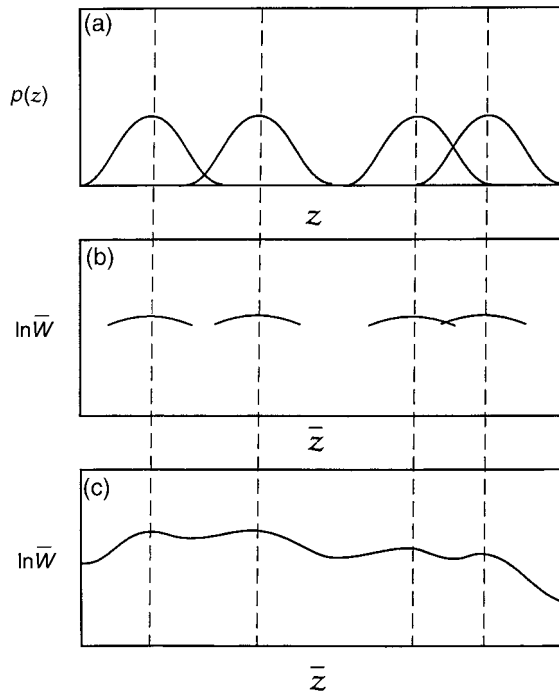


Figure 9. Local and global views of the adaptive landscape. (a) Four superimposed phenotypic distributions represent four different species. (b) In the local view of the adaptive landscape, the mean of each phenotypic distribution is situated at the optimum of a different adaptive landscape. The portions of these four landscapes in the vicinity of their optima are superimposed in the middle panel. (c) In the global view of the adaptive landscape, the phenotypic means are situated on different peaks of a single adaptive landscape.

situations in which fitness with respect to one trait depends upon the values of another trait. Such situations must be extremely common. They arise, for example, when traits function together in locomotion, foraging, sexual encounter, and all other fitness-related activities. Thus, the effect that a certain number of tail vertebrae has on crawling performance in snakes depends on how many body vertebrae an individual has. If a snake has more tail vertebrae, the optimal number of body vertebrae is higher than for a snake with fewer tail vertebrae (Arnold & Bennett, 1988). Put another way, snakes with the best crawling performance have a higher positive correlation in their vertebral numbers than the population at large. Such trait interactions affect the configuration and orientation of the individual selection surface and hence of the adaptive landscape. These aspects of the internal environment are encapsulated in the off-diagonal elements of the \mathbf{y} -matrix. In the snake example, the effect of the interaction is to produce a positively-oriented ridge in the surface that relates crawling speed to vertebral numbers. The

point is that the adaptive landscape is affected by trait interactions as well as by the ecological setting.

Critiques of the landscape idea

A variety of criticisms have been leveled at the adaptive landscape in the 70 years since its inception. Of these criticisms, the most fundamental is question of whether the landscape actually exists. The logical underpinnings of the landscape have been most thoroughly explored in the field of population genetics using one and two locus models. The earliest indications of problems were discovered in this arena by the originator of the concept, Sewall Wright. The main issue under discussion is whether evolution maximizes mean population fitness, that is, whether the population inevitably moves uphill on Wright's landscape and achieves a fitness maximum at equilibrium. Wright (1955, 1969) found that mean fitness was not maximized under certain forms of frequency-dependent selection, but in certain of these cases, another quantity could be defined that was maximized. Lewontin (1958) and Curtsinger (1984a, 1984b) searched for a quantity that would be maximized in a variety of problematic cases. They could show that such a quantity exists in some but not all cases. Ewens (1979) argued that fitness maximization depends on an assumption of linkage equilibrium and so is not a general evolutionary principle. This objection can be answered by assuming only weak selection and linkage (Kimura, 1965; Wright, 1969). From these analyses one can conclude that although Wright's landscape is a useful concept in many situations, there are circumstances in which its fundamental principles (increase in and maximization of mean fitness) do not apply. An alternative function may exist that is inevitably maximized during evolution, but so far no one has been able to find it.

The issue of fitness maximization has also been explored in the arena of Simpson-Lande landscapes. The maximization principle holds if the fitnesses of phenotypes are constant in models with single or multiple characters (Lande, 1976a, 1979). If selection is frequency-dependent, average fitness may not always increase in the population and the phenotypic mean may reach an equilibrium downslope from an optimum (Lande, 1976a, 1980a, 1981). In such cases, average fitness may not follow simple gradient dynamics, but it may still be useful to depict evolution on an adaptive landscape, for example, to visualize the effects of genetic covariance (Lande, 1980b).

Provine (1986) has criticized both Wright and Simpson–Lande landscapes, but on different grounds. Provine’s main complaint with Wright’s landscape is that it is often confused with an individual selection surface in which the axes are particular genotypic combinations (Provine, 1986, pp. 310–311). As Provine points out, such an individual selection surface is not a continuous function and so it cannot be the surface portrayed in Wright’s (1932) diagrams. With regard to Wright’s landscape (mean fitness as a function of gene frequency), Provine has no substantive criticism. Turning to the Simpson–Lande landscape, Provine’s main objection is that the evolutionary dynamics of the phenotypic mean are not formally related to an underlying theory of change in gene frequencies. A tractable theory for phenotypic evolution explicitly rooted in equations for genetic change at multiple loci is indeed a goal that has eluded theoreticians. The considerable progress that has been made in developing a useful evolutionary theory of phenotypes (Lande, 1988) was achieved by purposefully disconnecting that theory from population genetics and hence from its failure to achieve a polygenic extension. Whether one views this disconnection as an Achilles’ heel or an enabling tactic, depends on one’s outlook and priorities.

The Wrightian landscape is also at the center of a controversy over Wright’s shifting balance theory, but the issues of contention lose their force when applied to the Simpson–Lande landscape. The main issues of contention is whether populations become trapped on a suboptimal peaks and then overcome this condition through the joint agency of drift and interdemic selection (Whitlock & Phillips, 2000). The trapped situation arises on Wrightian landscapes because epistasis in fitness makes the landscape rugged (Whitlock et al., 1995). It is by no means clear that epistasis in fitness will play a comparable role on the Simpson–Lande landscape. In a highly polygenic world, the landscape of phenotypic traits is likely to be smooth. The prospect of becoming trapped is also exacerbated by the assumption that the landscape is constant through time. If the landscape ripples as its peak(s) move about, the population mean may work its way to the highest peak, even in the absence of genetic drift and interdemic selection.

Gould (1997) is dismayed by Dawkin’s (1996) *Mount Improbable* – a verbal portrayal of the Simpson–Lande landscape – because of what it leaves out. Gould prefers Lewontin’s metaphor of an environmental trampoline; “since organisms help to create

their own environment, adaptive peaks are built by interaction and undergo complex shifts as populations move in morphospace”. Gould is disappointed because Dawkin’s and Simpson’s landscapes leave out interaction between the organism and its environment, levels of selection, and other complexities that add richness to the discipline of evolutionary biology. Likewise, Eldredge and Cracraft (1980) and Eldredge (1999) object to the landscape concept because it leaves out selection at and above the level of species. None of these objections challenge the reality that we must approach the modeling of evolutionary processes in deliberate steps. The important point overlooked by all of the critiques just cited, as well as Dawkins (1996), is that the landscape concept is more than a metaphor. The landscape is a portrayal of a set of equations, not a bald invention. Those equations represent a growing set of models that capture an increasingly wider range of evolutionary possibilities. We may wish for models (and metaphors) that capture all possibilities now, but in the meantime the most tangible way to progress conceptually is to test and extend the models that we have.

What makes the adaptive landscape stable?

Under the landscape view of macroevolution the stability of the adaptive landscape seems an inescapable fact. The Bauplans that are often characteristic of genera and higher taxa can be understood as manifestations of a stable landscape. The cause of long-term stability of the landscape remains, however, an incompletely solved problem. Williams (1992) refers to it as a ‘desperation hypothesis’. The problem of stability is lessened if we remember that the adaptive landscape is not just an environmental phenomenon. To say that ‘the landscape is stable’ is not to say that ‘the environment is stable’. Organisms interact with their environment and some kinds of interactions can produce stability. Habitat selection (Partridge, 1978), for example, is a powerful behavioral mechanism that can compensate for environmental change and hence promote landscape stability. Trait interaction is another potential cause of stability. Traits that work together produce ridges, saddles and other topographic features of the adaptive landscape. It seems plausible that such features, arising from trait interactions, lend stability to the landscape. Nevertheless, landscape stability is an issue that needs more theoretical and empirical attention.

The theory of G-matrix evolution

The adaptive landscape provides the theoretical basis for a connection between microevolution and macroevolution, but to understand fully the flux of the landscape and the evolutionary response of a population to a changing landscape, we need to understand the genetic underpinnings of the multivariate phenotype. As noted above, important aspects of the genetics of the multivariate phenotype can be described statistically using the **G**-matrix. While most applications of the **G**-matrix assume that it remains relatively constant over evolutionary time, such an assumption may not always be valid. Unfortunately, the evolutionary dynamics of the **G**-matrix are not well understood. Despite more than two decades of effort, a dynamic analytical theory for the evolution of the **G**-matrix has not been produced, because the mathematical challenges have so far proven insurmountable. The problem has remained intractable because **G**-matrix stability depends on numerous factors, such as the number of loci affecting traits, the distribution of allelic effects at the loci, and the number of alleles per locus (Barton & Turelli, 1987; Turelli, 1988). One conclusion from existing models of the **G**-matrix is that analytical theory cannot guarantee **G**-matrix stability (Shaw et al., 1995), but the problem is so complex that existing theory cannot adequately describe the dynamics of the **G**-matrix over relevant periods of evolutionary time. Taken together, empirical and theoretical results indicate that the **G**-matrix may or may not be stable over multiple generations, leaving the question of **G**-matrix stability an unresolved issue. Future theoretical work involving both simulations and analytical models, coupled with careful empirical studies, may shed additional light on this important topic.

Summary

Is the ‘modern synthesis’ incomplete? Eldredge and Cracraft (1980) argue that microevolutionary processes cannot logically be extrapolated to explain macroevolutionary pattern. This argument seems to evaporate with the demonstration that among-taxa patterns of trait covariance can be predicted from models of microevolutionary process. Furthermore, the predictions can be compared against null models on a phylogeny. So long as that phylogeny includes higher taxa (e.g., genera) as well as populations and species, the extrapolation seems logically complete.

Turning to the other main complaint – that important pattern-producing processes operate above the level of populations – accommodation within the framework of adaptive landscapes seems possible. In Simpson’s (1944) visualization of adaptive zones, for example, lineage-specific extinction is diagrammed and linked to trait values. Continued construction of Simpson’s bridge, rather than demolition, may be the best path forward.

Our review exonerates Simpson’s vision of a landscape and suggests empirical and conceptual bridges between the often separate endeavors of studying micro- and macroevolution. Models of microevolutionary process that make predictions about macroevolutionary pattern provide the bridge between these endeavors. The key features of microevolutionary process include quantitative inheritance (genetic variances and covariances), effective population size, and configuration of the adaptive landscape (especially peak position and local curvature). Analyses of phenotypic variation within populations can characterize both key features of inheritance and local features of the landscape. Models that predict pattern from ecological processes are still poorly developed. In particular, the elusive concept of ecological opportunity deserves more theoretical and empirical attention. Thus, although the first generation of models provides many insights, much remains to be discovered.

Acknowledgements

We are grateful to Thomas Hansen, Andrew Hendry, Russell Lande, Emilia Martins, and two anonymous reviewers for helpful comments on the manuscript. This work was supported by a NSF research grant to S.J. Arnold and M.E. Pfrender and a NIH postdoctoral fellowship to A.G. Jones.

Appendix

The results that follow depend on a series of assumptions (Lande 1976a, 1979). The phenotypic distribution of traits, $p(z)$, is assumed to be (multivariate) normal, at least on some scale of measurement. The distribution of the breeding (additive genetic) values for the traits is assumed to be (multivariate) normal. This assumption of normally distributed breeding values does not necessarily imply that a very large

number of genes affect each trait. The assumption does require that more than a few genes affect each trait (so that the central limit theorem applies), and that no one gene explains the majority of genetic variance. Equations for the response of the phenotypic mean, \bar{z} , to selection allow any form for the individual fitness function (selection surface), $W(z)$, unless some special function is mentioned. Simple extrapolation of the selection response across more than one generation requires that phenotypic and additive genetic variances and covariances remain constant.

Matrices and vectors are shown in boldface. All matrices have n rows and n columns, where n is the number of traits. All vectors are column vectors with n elements. The superscript T denotes transpose. The superscript -1 denotes matrix inverse. For a brief review of matrix operations in the context of multivariate inheritance and selection see Arnold (1994).

(1) If selection acts on a single character, the change in the mean of that character, before selection, from one generation to the next is

$$\Delta \bar{z} = G\beta,$$

where G is the additive genetic variance of the character, and β is the directional selection gradient (Lande, 1976a, p. 317). This gradient, the slope of the adaptive landscape evaluated at the trait mean, is defined as

$$\beta = \frac{\partial \ln \bar{W}}{\partial \bar{z}}.$$

(2) The change in the natural log of average fitness in response to selection is approximately equal to generalized genetic distance. Because this distance is equal to or greater than zero, we can conclude that the effect of selection is to increase average fitness or leave it unchanged,

$$\Delta \ln \bar{W} \approx \Delta \bar{z}^T \mathbf{G}^{-1} \Delta \bar{z} \geq 0$$

(Lande, 1979, p. 406).

(3) The change in the multivariate mean, before selection, from one generation to the next is

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

where \mathbf{G} is the additive genetic variance–covariance matrix, and $\boldsymbol{\beta}$ is a vector of directional selection gradients,

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix},$$

whose n elements, corresponding to the n characters, are

$$\beta_i = \frac{\partial \ln \bar{W}}{\partial \bar{z}_i}$$

(Lande, 1979, p. 406).

(4) The curvature of the adaptive landscape in character dimensions z_i and z_j , evaluated at the trait mean, is

$$\gamma_{ij} - \beta_i \beta_j = \frac{\partial^2 \ln \bar{W}}{\partial \bar{z}_i \partial \bar{z}_j}$$

(Lande, 1979, p. 406; Phillips and Arnold, 1989, p. 1214), where γ_{ij} is an element in the $\boldsymbol{\gamma}$ -matrix. In the two character case, the $\boldsymbol{\gamma}$ -matrix is

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{bmatrix},$$

where γ_{11} and γ_{22} describe curvature (negative if curvature is downward, positive if curvature is upward) in trait dimensions z_1 and z_2 , respectively, and γ_{12} describes the orientation of the surface (positive if the surface tilts upward, negative if it tilts downward).

(5) The \mathbf{G} -matrix can be expressed in terms of its principal components (eigenvectors) and eigenvalues,

$$\mathbf{G} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^T,$$

where \mathbf{M} is an orthogonal matrix whose columns are the principal components of \mathbf{G} , and $\boldsymbol{\Lambda}$ is a matrix with the eigenvalues of \mathbf{G} on its diagonal and zeros elsewhere (Flury, 1988, Chap. 2). Abundant genetic variation in a particular direction in trait-space (principal component) is equivalent to a large eigenvalue for the corresponding principal component.

(6) For multiple traits, the among-population variance–covariance matrix for trait means after t generations is

$$\text{COV}[\bar{\mathbf{z}}] = \left(\frac{t}{N_e} \right) \bar{\mathbf{G}},$$

where N_e is effective population size and $\bar{\mathbf{G}}$ is the time average of \mathbf{G} (Lande, 1979, pp. 408–409).

(7) For multiple traits after t generations,

$$\text{COV}[\bar{\mathbf{z}}] = t\mathbf{G}\text{COV}[\boldsymbol{\beta}]\mathbf{G},$$

where $\text{COV}[\boldsymbol{\beta}]$ is the among-population variance–covariance matrix for $\boldsymbol{\beta}$, the directional selection gradient, and

$$\text{COV}[\boldsymbol{\beta}] = \mathbf{P}^{-1}\mathbf{s}\mathbf{P}^{-1},$$

where \mathbf{P} is the variance–covariance matrix for the traits before selection, and \mathbf{s} is the column vector of

directional selection differentials (Felsenstein, 1988, pp. 451–452; Zeng, 1988, p. 370). For a single trait, the among-population variance in trait mean after t generations is

$$\text{VAR}(\bar{z}) = t(G^2/P^2)\text{VAR}(s),$$

where G is additive genetic variance, P is phenotypic variance before selection and $\text{Var}(s)$ is the among-population variance in the directional selection differential for the trait.

(8) For multiple traits,

$$\text{COV}[\bar{\mathbf{z}}] = t[\mathbf{G}/N_e + \mathbf{G}\text{COV}[\boldsymbol{\beta}]\mathbf{G}]$$

(Hansen & Martins, 1996, p. 1409).

(9) For multiple traits, the restraining force is

$$\mathbf{G}[\mathbf{P} - \boldsymbol{\gamma}^{-1}] \approx -\mathbf{G}\boldsymbol{\gamma},$$

assuming weak selection (Hansen & Martins, 1996, p. 1410). Note that if selection is weak, so that $\beta \approx 0$ and $\boldsymbol{\omega} \gg \mathbf{P}$, where $\boldsymbol{\omega}$ is a matrix whose elements describe a Gaussian individual selection surface (elements of $\boldsymbol{\omega}$ analogous to variances and covariances), then

$$\boldsymbol{\omega} \approx -\boldsymbol{\gamma}^{-1}$$

(Lande, 1979, pp. 406–407). Hansen and Martin (1996) use \mathbf{W} to represent $\boldsymbol{\omega}^{-1}$. In the univariate case, ω^2 is sometimes used to represent the ‘variance’ of a Gaussian selection surface (e.g., Lande, 1976), in which case $\omega^2 \approx -1/\gamma$.

(10) For multiple traits after t generations,

$$\text{COV}[\bar{\mathbf{z}}] = \mathbf{Q}(t)\hat{\mathbf{V}}\mathbf{Q}^T(t),$$

where

$$\mathbf{Q}(t) = \exp(t\mathbf{G}\boldsymbol{\gamma})$$

and

$$\hat{\mathbf{V}} = -(2N_e\boldsymbol{\gamma})^{-1}$$

(Hansen & Martins 1996, p. 1410). For a single trait after t generations,

$$\text{Var}(\bar{z}) = -(2N_e\gamma)^{-1} \exp(2tG\gamma).$$

(11) The expected time until the trait mean shifts to the second adaptive peak is

$$T \propto (\bar{W}_p \bar{W}_v)^{2N_e},$$

where \bar{W}_p represents the height of the original peak and \bar{W}_v represents the height of the valley between the two peaks (Lande, 1986, p. 345).

(12) For a single trait, the expected change in the trait mean is

$$\Delta \bar{z} = -G\gamma d,$$

where d is the distance of the trait mean from the optimum (Lande & Shannon, 1996, p. 435).

(13) For multiple traits in large populations after t generations,

$$\text{COV}[\bar{\mathbf{z}}] = \mathbf{Q}(t)\text{COV}[\boldsymbol{\theta}]\mathbf{Q}^T(t),$$

where $\text{COV}[\boldsymbol{\theta}]$ is the among-population variance–covariance matrix for the optima of the adaptive landscapes (Hansen & Martins, 1996, p. 1411). For a single trait after t generations,

$$\text{Var}(\bar{z}) = \text{Var}(\theta) \exp(2tG\gamma),$$

where $\text{Var}(\theta)$ is the among-population variance in the optima of the adaptive landscapes.

(14) Let the individual selection surface be a Gaussian surface with optimum $\boldsymbol{\theta}$,

$$W(\mathbf{z}) = \exp\left[-\frac{1}{2}(\mathbf{z} - \boldsymbol{\theta})^T \boldsymbol{\omega}^{-1}(\mathbf{z} - \boldsymbol{\theta})\right].$$

Suppose the optimum of the surface moves as a normally distributed random variable with mean 0 and a variance–covariance matrix that is proportional to $\boldsymbol{\omega}$. Then, if selection is weak, so that $\boldsymbol{\beta} \approx \mathbf{0}$ and $\boldsymbol{\omega} \gg \mathbf{P}$,

$$\text{COV}(\boldsymbol{\theta}) = k\boldsymbol{\omega} = -k\boldsymbol{\gamma}^{-1},$$

where k is a constant of proportionality (a scalar). Then, using (13), the among-population variance–covariance matrix for trait means after t generations is

$$\text{COV}[\bar{\mathbf{z}}] = \mathbf{Q}(t)(k\boldsymbol{\omega})\mathbf{Q}^T(t) \approx \mathbf{Q}(t)(-k\boldsymbol{\gamma}^{-1})\mathbf{Q}^T(t).$$

For a single trait after t generations,

$$\begin{aligned} \text{Var}(\bar{z}) &= (k\omega) \exp(-2tG/\omega) \\ &\approx (-k/\gamma) \exp(2tG\gamma). \end{aligned}$$

(15) The evolutionary load caused by the deviation of the trait mean from its optimum after t generations is

$$(\gamma/2)E[\bar{z}(t) - \theta(t)]^2,$$

where the second term represents the expected value of the squared deviation of the trait mean from its optimum after t generations (Lande & Shannon, 1996, p. 435).

(16) The directional selection differential is the change in trait mean within a generation arising from directional selection

$$\mathbf{s} = [\bar{\mathbf{z}}^* - \bar{\mathbf{z}}] = \mathbf{P}\boldsymbol{\beta},$$

where $\bar{\mathbf{z}}^*$ is the trait mean after selection and $\bar{\mathbf{z}}$ is the trait mean before selection (Lande & Arnold, 1983, p. 1213). The shift in the mean of trait z_1 due to directional selection is

$$s_1 = (\bar{z}_1^* - \bar{z}_1) = P_{11}\beta_1 + P_{12}\beta_2 + \cdots + P_{1n}\beta_n,$$

where $P_{11}\beta_1$ is the portion of the shift due to selection on trait z_1 , $P_{12}\beta_2$ is the portion of the shift due to selection on trait z_2 , and $P_{1n}\beta_n$ is the portion of the shift due to selection on trait z_n .

(17) The nonlinear selection differential is the change in trait variance and covariance within a generation arising from nonlinear selection

$$\mathbf{C} = \mathbf{P}^* - \mathbf{P} + \mathbf{s}\mathbf{s}^T = \mathbf{P}\boldsymbol{\gamma}\mathbf{P},$$

where \mathbf{P}^* is the trait variance–covariance matrix after selection (Lande & Arnold, 1983, p. 1216). In the case of two traits under selection, the change in the variance of trait z_1 due to nonlinear selection is

$$\begin{aligned} C_{11} &= P_{11}^* - P_{11} + s^2 \\ &= P_{11}^2\gamma_{11} + 2P_{11}P_{12}\gamma_{12} + P_{12}^2\gamma_{22}, \end{aligned}$$

where the s_1^2 term corrects for the decrease in variance arising from directional selection. The $P_{11}^2\gamma_{11}$ term describes the change in variance due to nonlinear selection on trait z_1 (γ_{11} is negative in the case of stabilizing selection and positive in the case of disruptive selection). The $P_{11}P_{12}\gamma_{12}$ term describes the change in variance of trait z_1 due to correlational selection on traits z_1 and z_2 . The $P_{12}^2\gamma_{22}$ term describes the change in variance of trait z_1 due to nonlinear selection on trait z_2 .

(18) The individual selection surface can often be approximated by a quadratic surface, which in the case of two traits, is

$$\begin{aligned} W(z) &= \alpha + \beta_1 z_1 + \beta_2 z_2 + \frac{1}{2}\gamma_{11} z_1^2 \\ &\quad + \frac{1}{2}\gamma_{22} z_2^2 + \gamma_{12} z_1 z_2, \end{aligned}$$

where α is a constant, z_1 and z_2 are the trait values for the two traits (standardized so that their means = 0), β_1 and β_2 are the directional selection gradients for the two traits, γ_{11} and γ_{22} are the nonlinear selection gradients describing stabilizing (or disruptive) selection, and γ_{12} is the nonlinear selection gradient describing correlational selection. This quadratic surface can be estimated from data using the quadratic regression model

$$\begin{aligned} w &= \alpha + \beta_1 z_1 + \beta_2 z_2 + \frac{1}{2}\gamma_{11} z_1^2 + \frac{1}{2}\gamma_{22} z_2^2 \\ &\quad + \gamma_{12} z_1 z_2 + \varepsilon, \end{aligned}$$

where w is relative fitness (standardized so that its mean = 1), α is a constant, and ε is an error term (Lande & Arnold, 1983, p. 1217).

(19) The directional selection gradient is the slope of the individual selection surface, $\partial W(\mathbf{z})/\partial \mathbf{z}$, weighted by the phenotypic trait distribution before selection, $p(\mathbf{z})$;

$$\boldsymbol{\beta} = \int p(\mathbf{z}) \frac{\partial W(\mathbf{z})}{\partial \mathbf{z}} d\mathbf{z}$$

(Lande & Arnold, 1983, p. 1213).

(20) The nonlinear selection gradient is the curvature of the individual selection surface, $\partial^2 W(\mathbf{z})/\partial^2 \mathbf{z}$, weighted by $p(\mathbf{z})$;

$$\boldsymbol{\gamma} = \int p(\mathbf{z}) \frac{\partial^2 W(\mathbf{z})}{\partial^2 \mathbf{z}} d\mathbf{z}$$

(Lande & Arnold, 1983, p. 1216).

(21) The net selection gradient corresponding to the phenotypic divergence of two sister taxa is the sum of over generations of the directional selection gradients, $\boldsymbol{\beta}_t$, that have acted during the period of divergence. If \mathbf{G} remains constant during the period of divergence, the net selection gradient can be estimated by the formula,

$$\boldsymbol{\beta}_{\text{net}} = \mathbf{G}^{-1}[\bar{\mathbf{z}}_a - \bar{\mathbf{z}}_b],$$

where $\bar{\mathbf{z}}_a$ and $\bar{\mathbf{z}}_b$ are the trait means of two sister taxa (Lande, 1979, p. 407). If \mathbf{G} varies during the period of divergence, the net divergence in means is a function of the covariance between \mathbf{G} and $\boldsymbol{\beta}$, $\text{COV}(\mathbf{G}_t, \boldsymbol{\beta}_t)$, as well as the time average of \mathbf{G} and the sum of $\boldsymbol{\beta}_t$ (Turelli, 1988). Although Turelli (1988) argued that $\text{COV}(\mathbf{G}_t, \boldsymbol{\beta}_t)$ might be large, it is not clear that this is so. Selection in generation t affects \mathbf{G} in generation $t + 1$. Thus, we might expect a nonzero $\text{COV}(\mathbf{G}_{t+1}, \boldsymbol{\beta}_t)$, if $\boldsymbol{\beta}_t$ varies. Nonzero $\text{COV}(\mathbf{G}_t, \boldsymbol{\beta}_t)$ would depend on serial autocorrelation in $\boldsymbol{\beta}_t$. This autocorrelation obviously depends on how the adaptive landscape varies through time, which is an unresolved empirical issue.

References

- Arnold, S.J., 1988. Quantitative genetics and selection in natural populations: microevolution of vertebral numbers in the garter snake, *Thamnophis elegans*, pp. 619–636 in Proc. 2nd Intern. Conf. Quantitative Genetics, edited by B.S. Weir, E.J. Eisen, M.M. Goodman & G. Namkoong. Sinauer, Sunderland, M.A.
- Arnold, S.J., 1992. Constraints on phenotypic evolution. Am. Nat. 140: S85–S107.

- Arnold, S.J., 1994. Multivariate inheritance and evolution: a review of the concepts, pp. 17–48 in *Quantitative Genetic Studies of the Evolution of Behavior*, edited by C.R.P. Boake. University of Chicago Press, Chicago.
- Arnold, S.J. & A.F. Bennett, 1988. Behavioral variation in natural populations. V. Morphological correlates of locomotion in the garter snake *Thamnophis radix*. *Biol. J. Linn. Soc.* 34: 175–190.
- Arnold, S.J. & P.C. Phillips, 1999. Hierarchical comparison of genetic variance–covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. *Evolution* 53: 1516–1527.
- Badyaev, A.V. & G.E. Hill, 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence on morphological covariance structure. *Evolution* 54: 1784–1794.
- Barton, N.H. & M. Turelli, 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res.* 49: 157–173.
- Brodie, E.D. III, 1992. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. *Evolution* 47: 1284–1298.
- Brodie, E.D. III, A.J. Moore & F.J. Janzen, 1995. Visualizing and quantifying natural selection. *Trends Ecol. Evol.* 10: 313–318.
- Brown, J.S. & T.L. Vincent, 1992. Organization of predator-prey communities as an evolutionary game. *Evolution* 46: 1269–1283.
- Bull, J.J., 1987. Evolution of phenotypic variance. *Evolution* 41: 303–315.
- Charlesworth, B., 1993a. The evolution of sex and recombination in a varying environment. *J. Hered.* 84: 345–350.
- Charlesworth, B., 1993b. Directional selection and the evolution of sex and recombination. *Genet. Res.* 61: 205–224.
- Charlesworth, B., R. Lande & M. Slatkin, 1982. A neo-Darwinian commentary on macroevolution. *Evolution* 36: 474–498.
- Cheverud, J.M., 1984. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* 110: 155–171.
- Curtsinger, J.W., 1984a. Evolutionary landscapes for complex selection. *Evolution* 38: 359–367.
- Curtsinger, J.W., 1984b. Evolutionary principles for polynomial models of frequency-dependent selection. *Proc. Natl. Acad. Sci.* 81: 2840–2842.
- Dawkins, R., 1996. *Climbing Mount Improbable*. W.W. Norton, New York.
- Dudley, S.A., 1996. The response to differing selection on plant physiological traits: evidence for local adaptation. *Evolution* 50: 103–110.
- Eldredge, N., 1999. *The Pattern of Evolution*. W.H. Freeman and Co., New York.
- Eldredge, N. & J. Cracraft, 1980. *Phylogenetic Patterns and the Evolutionary Process*. Columbia University Press, New York, N.Y.
- Emerson, S.B. & S.J. Arnold, 1989. Intra- and interspecific relationships between morphology, performance, and fitness, pp. 295–314 in *Complex Organismal Functions: Integration and Evolution*, edited by D.B. Wake & G. Roth. Wiley, Chichester, U.K.
- Endler, J.A., 1986. *Natural Selection in the Wild*. Princeton University Press, Princeton, N.J.
- Ewens, W.J., 1979. *Mathematical Population Genetics*. Springer-Verlag, New York.
- Feder, J.F., 1998. The apple maggot fly, *Rhagoletis pomonella*: flies in the face of conventional wisdom, pp. 130–144 in *Endless Forms: Species and Speciation*, edited by D.J. Howard & S.H. Berlocher. Oxford University Press, New York.
- Felsenstein, J., 1973. Maximum likelihood estimation of evolutionary trees from continuous characters. *Am. J. Hum. Genet.* 25: 471–492.
- Felsenstein, J., 1979. Excursions along the interface between disruptive and stabilizing selection. *Genetics* 93: 773–795.
- Felsenstein, J., 1985. Phylogenies and the comparative method. *Am. Nat.* 125: 1–15.
- Felsenstein, J., 1988. Phylogenies and quantitative characters. *Ann. Rev. Ecol. Syst.* 19: 445–471.
- Flury, B., 1988. *Common Principal Components and Related Multivariate Models*. Wiley, New York, N.Y.
- Garcia-Ramos, G. & M. Kirkpatrick, 1997. Genetic models of adaptation and gene flow in peripheral populations. *Evolution* 51: 21–28.
- Gavrilets, S., 1997. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* 12: 307–312.
- Gould, S.J., 1997. Self-help for a hedgehog stuck on a molehill. *Evolution* 51: 1020–1023.
- Gould, S.J. & N. Eldredge, 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiol.* 3: 115–151.
- Hansen, T.F. & E.P. Martins, 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50: 1404–1417.
- Hansen, T.F., 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51: 1341–1351.
- Hendry, A.P. & M.T. Kinnison, 1999. The pace of modern life: measuring rates of microevolution. *Evolution* 53: 1637–1653.
- Hendry, A.P., J.K. Wenburg, P. Bentzen, E. C. Volk & T. P. Quinn, 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290: 516–518.
- Hendry, A.P., T. Day & E. B. Taylor, 2001. Population mixing and the adaptive divergence of quantitative characters in discrete populations: a theoretical framework for empirical tests. *Evolution* 55: 459–466.
- Janzen, F.J. & H.S. Stern, 1998. Logistic regression for empirical studies of multivariate selection. *Evolution* 52: 1564–1571.
- Karn, M.L. & L.S. Penrose, 1951. Birth weight and gestation time in relation to maternal age, parity, and infant survival. *Ann. Eugenics* 16: 147–164.
- Kimura, M., 1965. Attainment of quasi-linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* 52: 875–890.
- Kingsolver, J.G., H.E. Hoekstra, J.M. Hoekstra, D. Berrigan, S.N. Vignieri, C.E. Hill, A. Hoang, P. Gilbert & P. Beerli, 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157: 245–261.
- Kirkpatrick, M., 1982. Quantum evolution and punctuated equilibrium in continuous genetic characters. *Am. Nat.* 119: 833–848.
- Kurtén, B., 1959. Rates of evolution in fossil mammals. *Cold Spring Harbor Symp. Quant. Biol.* 24: 205–215.
- Lande, R., 1976a. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314–334.
- Lande, R., 1976b. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* 26: 221–235.
- Lande, R., 1979. Quantitative genetic analysis of multivariate evolution, applied to brain-body size allometry. *Evolution* 33: 402–416.
- Lande, R., 1980a. Microevolution in relation to macroevolution. *Paleobiol.* 6: 233–238.
- Lande, R., 1980b. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34: 292–305.
- Lande, R., 1980c. The genetic covariance between characters maintained by pleiotropic mutation. *Genetics* 94: 203–215.

- Lande, R., 1981. The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* 99: 541–553.
- Lande, R., 1986. The dynamics of peak shifts and the pattern of morphological evolution. *Paleobiol.* 12: 343–354.
- Lande, R., 1988. Genetics and demography in biological conservation. *Science* 241: 1455–1460.
- Lande, R. & S.J. Arnold, 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
- Lande, R. & S. Shannon, 1996. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* 50: 434–437.
- Lewontin, R.C., 1958. A general method for investigating the equilibrium of gene frequencies in a population. *Genetics* 43: 419–434.
- Lynch, M., 1990. The rate of morphological evolution in mammals from the standpoint of neutral expectation. *Am. Nat.* 136: 727–741.
- Lynch, M. & R. Lande, 1993. Evolution and extinction in response to environmental change, pp. 234–250 in *Biotic Interactions and Global Change*, edited by P. Kareiva, J.G. Kingsolver & R.B. Huey. Sinauer, Sunderland, MA.
- Martins, E.P. & T.F. Hansen, 1996. A microevolutionary link between phylogenies and comparative data, pp. 273–288 in *New Uses of New Phylogenies*, edited by P.H. Harvey, A.J. Leigh Brown, J. Maynard Smith & S. Nee. Oxford University Press, Oxford, U.K.
- Mather, K., 1941. Variation and selection of polygenic characters. *J. Genetics* 41: 159–193.
- Partridge, L., 1978. Habitat selection, pp. 351–376 in *Behavioral Ecology*, edited by J.R. Krebs & N.B. Davies. Blackwell Scientific, Oxford, England.
- Pfrender, M., 1998. Evolutionary dynamics of molecular and quantitative genetic variation in ephemeral pond populations of *Daphnia pulex*. Thesis Dissertation, University of Oregon, Eugene, Oregon.
- Phillips, P.C. & S.J. Arnold, 1989. Visualizing multivariate selection. *Evolution* 43: 1209–1222.
- Phillips, P.C. & S.J. Arnold, 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution* 53: 1506–1515.
- Phillips, P.C., M.C. Whitlock & K. Fowler, 2001. Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. *Genetics* 158: 1137–1145.
- Price, T.D., P.R. Grant & P.T. Boag, 1984. Genetic changes in the morphological differentiation of Darwin's ground finches, pp. 49–66 in *Population Biology and Evolution*, edited by K. Wohrman & V. Loeschke. Springer-Verlag, Berlin.
- Provine, W.B., 1986. *Sewall Wright and Evolutionary Biology*. University of Chicago Press, Chicago, I.L.
- Rensch, B., 1959. *Evolution Above the Species Level*. Wiley, New York, N.Y.
- Reznick, D.N. & C.K. Ghalambor, 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112–113: 183–198.
- Reznick, D.N., F.H. Shaw, F.H. Rodd & R.G. Shaw, 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* 275: 1934–1937.
- Rieseberg, L.H., M.A. Archer & R.K. Wayne, 1999. Transgressive segregation, adaptation and speciation. *Heredity* 83: 363–372.
- Rice, S.H., 1998. The evolution of canalization and the breaking of von Baer's laws: modeling the evolution of development with epistasis. *Evolution* 52: 647–656.
- Roff, D.A., 2000. The evolution of the **G** matrix: selection or drift? *Heredity* 84: 135–142.
- Rundle, H.D. & M.C. Whitlock, 2001. A genetic interpretation of ecologically dependent isolation. *Evolution* 55: 198–201.
- Schluter, D., 1984. Morphological and phylogenetic relations among the Darwin's finches. *Evolution* 38: 921–930.
- Schluter, D., 1988. Estimating the form of natural selection on a quantitative trait. *Evolution* 42: 849–861.
- Schluter, D., 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50: 1766–1774.
- Schluter, D. & D. Nychka, 1994. Exploring fitness surfaces. *Am. Nat.* 143: 597–616.
- Schluter, D., 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford, England.
- Schmalhausen, I.I., 1949. *Factors of Evolution, The Theory of Stabilizing Selection*. University of Chicago Press, Chicago, I.L.
- Shaw, F.H., R.G. Shaw, G.S. Wilkinson & M.H. Turelli, 1995. Changes in genetic variances and covariances: G whiz! *Evolution* 49: 1260–1267.
- Simpson, G.G., 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York, N.Y.
- Simpson, G.G., 1953. *The Major Features of Evolution*. Columbia University Press, New York.
- Sinervo, B., P. Doughty, R.B. Huey & K. Zamudio, 1992. Allometric engineering: a causal analysis of natural selection on offspring size. *Science* 258: 1927–1930.
- Slatkin, M. & R. Lande, 1976. Niche width in a fluctuating environment – density independent model. *Am. Nat.* 110: 31–55.
- Stanley, S.M., 1979. *Macroevolution: Pattern and Process*. W. H. Freeman and Co., San Francisco.
- Steppan, S.J., 1997a. Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. *Evolution* 51: 571–586.
- Steppan, S.J., 1997b. Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. *Evolution* 51: 587–594.
- Svensson, E. & B. Sinervo, 2000. Experimental excursions on adaptive landscapes: density-dependent selection on egg size. *Evolution* 54: 1396–1403.
- Thompson, J.N., 1998. Rapid evolution as an ecological process. *Trends Ecol. Evol.* 13: 329–332.
- Turelli, M., 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution* 42: 1342–1347.
- Via, S., 1991. The genetic structure of host plant adaptation in a spatial patchwork: demographic variability among reciprocally transplanted pea aphid clones. *Evolution* 45: 827–852.
- Vrba, E.S., 1983. Macroevolutionary trends: new perspectives on the roles of adaptation and incidental effect. *Science* 221: 387–389.
- Waddington, C.H., 1957. *The Strategy of the Genes*. MacMillan, New York, N.Y.
- Williams, G.C., 1992. *Natural Selection: Domains, Levels, and Challenges*. Oxford University Press, Oxford, U.K.
- Willis, J.H., J.A. Coyne & M. Kirkpatrick, 1991. Can one predict the evolution of quantitative characters without genetics? *Evolution* 45: 441–444.
- Whitlock, M.C., 1995. Variance-induced peak shifts. *Evolution* 49: 252–259.
- Whitlock, M.C. & P. C. Phillips, 2000. The exquisite corpse: a shifting view of the shifting balance. *Trends Ecol. Evol.* 15: 347–348.

- Whitlock, M.C., P.C. Phillips, F.B.-G. Moore & S.J. Tonsor, 1995. Multiple fitness peaks and epistasis. *Ann. Rev. Ecol. Syst.* 26: 601–629.
- Wright, S., 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.
- Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proc. VI Inter. Congr. Genet.* 1: 356–366.
- Wright, S., 1940. The statistical consequences of Mendelian heredity in relation to speciation, pp. 161–183 in *The New Systematics*, edited by J.S. Huxley. Clarendon Press, Oxford.
- Wright, S., 1945. Tempo and mode in evolution: a critical review. *Ecology* 26: 415–419.
- Wright, S., 1955. Classification of the factors of evolution. *Cold Spring Harbor Symp. Quant. Biol.* 20: 16–24.
- Wright, S., 1968. *Evolution and the Genetics of Populations*, Vol. 1. Genetic and Biometric Foundations. University of Chicago Press, Chicago, I.L.
- Wright, S., 1969. *Evolution and the Genetics of Populations*, Vol. 2. The Theory of Gene Frequencies. University of Chicago Press, Chicago, I.L.
- Wright, S., 1982. Character change, speciation, and the higher taxa. *Evolution* 36: 427–443.
- Zeng, Z.-B., 1988. Long-term correlated response, interpopulation covariance, and interspecific allometry. *Evolution* 42: 363–374.