EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 52 **June 1998** **No. 3**

Evolution, 52(3), 1998, pp. 647–656

THE EVOLUTION OF CANALIZATION AND THE BREAKING OF VON BAER'S LAWS: MODELING THE EVOLUTION OF DEVELOPMENT WITH EPISTASIS

SEAN H. RICE

Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520

E-mail: sean.rice@yale.edu

Abstract.—Evolution can change the developmental processes underlying a character without changing the average expression of the character itself. This sort of change must occur in both the evolution of canalization, in which a character becomes increasingly buffered against genetic or developmental variation, and in the phenomenon of closely related species that show similar adult phenotypes but different underlying developmental patterns. To study such phenomena, I develop a model that follows evolution on a surface representing adult phenotype as a function of underlying developmental characters. A contour on such a "phenotype landscape" is a set of states of developmental characters that produce the same adult phenotype. Epistasis induces curvature of this surface, and degree of canalization is represented by the slope along a contour. I first discuss the geometric properties of phenotype landscapes, relating epistasis to canalization. I then impose a fitness function on the phenotype and model evolution of developmental characters as a function of the fitness function and the local geometry of the surface. This model shows how canalization evolves as a population approaches an optimum phenotype. It further shows that under some circumstances, "decanalization" can occur, in which the expression of adult phenotype becomes increasingly sensitive to developmental variation. This process can cause very similar populations to diverge from one another developmentally even when their adult phenotypes experience identical selection regimes.

Key words.—Canalization, epistasis, evolution of development, phenotype landscape, quantitative genetics.

Received January 17, 1997. Accepted February 18, 1998.

The observation that embryos of related species tend to resemble one another more than do the adults that develop from them is one of the longest standing generalizations about comparative development. Thirty years before *On the Origin of Species* (Darwin 1859), K. E. von Baer (1828) codified the idea of embryonic conservatism in the third and fourth of his four laws of development. von Baer's third law states that the embryo of a particular type of animal diverges from that of other types as it develops, rather than passing through the same stages as others. The fourth law concludes from this that the embryos of "higher" animals will resemble the embryos, but not the adults, of other species (for a translation see Gilbert 1991).

Darwin (1859) gave this pattern an evolutionary interpretation, arguing "that certain organs in the individual, which when mature become widely different and serve for different purposes, are in the embryo exactly alike" was evidence in support of the idea of descent with modification. More recently, the idea that early developmental characters are evolutionarily conservative arises, implicitly or explicitly, in discussions of homology (Hall 1992), macroevolution (Levinton 1988), and the value of characters for phylogenetic inference (Kluge and Strauss 1985).

The continuing popularity of von Baer's third and fourth laws is well justified: a general pattern of embryonic conservatism is the norm in all major groups of animals. There are circumstances, though, in which developmental processes undergo substantial evolution that is not always manifest in adult phenotype—producing similar species that differ in development but not in adult morphology (Raff 1996).

In some such cases, divergence can be attributed to selection acting directly on early developmental stages associated with life-history changes such as a shift from indirect to direct development (Wray 1994). In other cases, though, we see changes in developmental stages that do not appear to be under direct selection and continue to produce the same adult morphology. For example, in early gastropod embryos, determination of the *D* quadrant (which includes cells that produce much of the shell gland and all of the mesoderm) is controlled by cytoplasmic localization at the second cleavage event in some species and by inductive interactions between fifth and sixth cleavage in others (Freeman and Lundelius 1992).

Similar divergence can occur later in development, when tissue interactions are determining patterns of differentiation.

For example, attempts to identify "the" mechanism of lens formation in the vertebrate eye have been hindered by the fact that the significance of the optic cup in inducing lens formation varies between different frog species, with the optic cup being necessary in some species but not in others (Jacobson and Sater 1988).

This sort of developmental divergence can appear even among populations experiencing the same selection regime. For example, Rutledge et al. (1974) found that replicate lines of inbred mice responded to selection for increased tail length in different ways: one increasing the number of vertebrae and the other increasing the length of each vertebra.

Similar divergence under uniform selection occurs at the genetic level (Cohan and Hoffmann 1989). For example, Cohan (1984) found that under uniform selection on a previously unselected trait in *Drosophila,* similar lines diverged genetically, and this divergence was as great between lines drawn from the same population as between lines drawn from different populations.

It is thus possible for developmental processes to evolve independently of the characters that they build. One case in which this must happen is the evolution of canalization, defined here as the degree to which a phenotypic character is buffered against variation in the underlying developmental processes that construct it. (This definition is equivalent to "genetic canalization" as used by Stearns and Kawecki [1994] and Wagner et al. [1997], who make the important distinction between genetic and environmental canalization. It is also consistent with Waddington's [1957] use of the term though without the emphasis on discreetness of developmental pathways.)

Canalization involves restructuring a developmental process so as to reduce the variance in the phenotype that results from that process without changing the mean value of that phenotype. This is only possible if developmental characters interact in nonadditive, or epistatic, ways.

In this paper, I will extend the term "epistasis" to refer to nonadditive interactions between developmental characters, as well as between loci. Extending this definition to developmental characters simply recognizes that nonadditive interactions between loci are often a result of interactions between developmental characters that those loci contribute to. For polygenic characters, developmental epistasis always implies genetic epistasis. Note that by this definition, epistasis is a property of an organism, not a population (as it often is when defined as a statistical deviation from additivity). It is thus similar to the idea of "physiological epistasis" used by Cheverud and Routman (1996).

Making sense of groups that break von Baer's laws and understanding the evolution of canalization requires that we model both the evolution of a phenotypic character and the nonadditive developmental processes that construct that character. One way to do this is to model evolution on a phenotype landscape. Here, I represent various values of a phenotype as a surface over a space of underlying developmental or genetic factors. Studying evolution on such a surface allows us to simultaneously watch change in the developmental factors and the phenotype that they build. I will first describe the geometry of phenotype landscapes, which captures the nature of developmental interactions and allows us to more carefully define such concepts as "canalization." I then bring selection into the picture and derive the equations of evolution in terms of the way selection acts on phenotype and the way phenotype is related to underlying development.

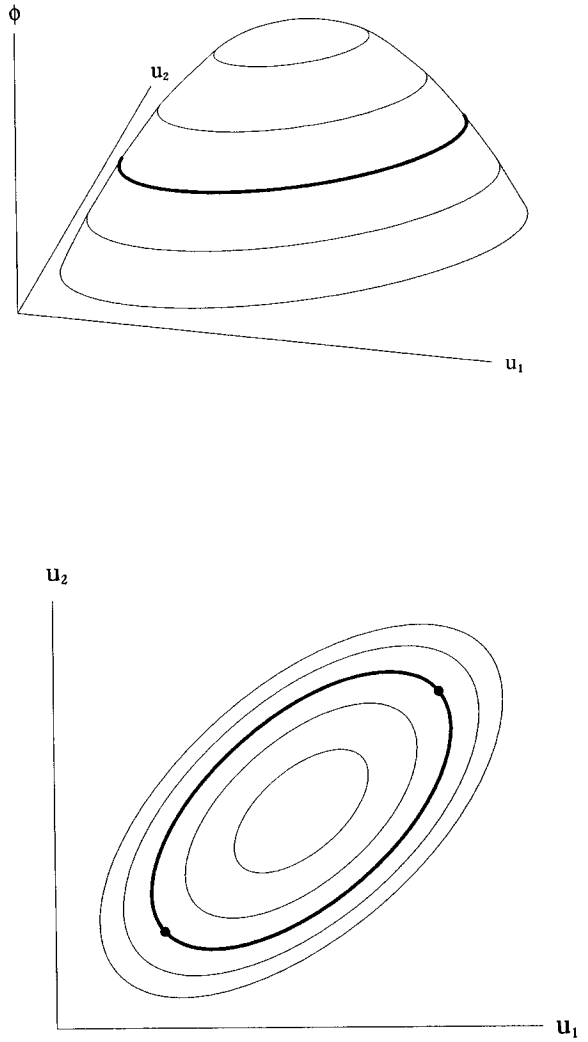

FIG. 1. A phenotype landscape over a space of two underlying developmental/genetic characters. The contours of this landscape represent sets of equal phenotype. The two dots in the second figure are points of maximum canalization along the bold phentoype contour.

GEOMETRY OF THE PHENOTYPE LANDSCAPE

Consider a phenotype, ϕ, which is a function of two underlying factors, u_1 and u_2. These underlying factors may represent quantitative genetic characters in their own right (Slatkin 1987) or continuous measures of the expression of some gene product. Plotting ϕ as a function of u_1 and u_2 yields a surface, the contours of which represent different phenotypic values. Figure 1 shows an example of such a plot. In general, if there are n underlying factors, then the surface is n dimensional and a particular phenotypic value is represented by an $n - 1$ dimensional contour. Selection may

cause a population to move up or down such a landscape or along a contour.

Note that this is different from an "adaptive landscape" over a space of phenotypic characters (Simpson 1944; Lande and Arnold 1983). The phenotype landscape is a property of an individual and its shape describes the nature of developmental interactions. A peak on a phenotype landscape represents the extreme value of a phenotypic character that can be achieved by local changes in development; it thus can be thought of as a constraint, but in general will not be an optimum.

The phenotype landscape is also different from the "epigenetic landscape" of Waddington (1957). On Waddington's surface, height was an abstract measure of a developmental potential function, such that the system moved downslope during development. On a phenotype landscape, height is simply a measure of some phenotypic character.

If all factors contribute additively to the phenotype, such that the consequence of changing a factor is not a function of its own state or that of other factors, then the phenotype surface is a plane. This is the tacit view of many quantitative genetic models. Interactions between the underlying factors (epistasis) or dependence of the effects of changing a factor on its current state cause the surface to curve.

At a given point, the slope of the surface captures the degree to which a given amount of variation in the underlying developmental factors translates into variation in the resulting phenotype. Slope thus provides a measure of the degree of canalization. In the purely additive case, slope is the same everywhere, so increased canalization cannot evolve. If the surface curves, though, then the slope is likely to be lower at some points along a contour than at others. Points of minimum slope along a contour are thus points of maximum canalization for that phenotypic value. Such points are properties of the phenotype landscape alone, and identifying them mathematically allows us to identify the important geometric properties of the phenotype landscape.

The most important terms for discussing the geometry of $\phi(u_1, u_2)$ are the phenotype gradient, $\nabla\phi$, defined as

$$\nabla\phi = \begin{bmatrix} \frac{\partial\phi}{\partial u_1} \\ \frac{\partial\phi}{\partial u_2} \end{bmatrix}, \tag{1}$$

and the epistasis matrix, E, defined as

$$E = \begin{bmatrix} \frac{\partial^2\phi}{\partial u_1^2} & \frac{\partial^2\phi}{\partial u_1 \partial u_2} \\ \frac{\partial^2\phi}{\partial u_1 \partial u_2} & \frac{\partial^2\phi}{\partial u_2^2} \end{bmatrix}. \tag{2}$$

The phenotype gradient is a vector that points in the direction of maximum slope on the surface. The gradient at a particular point is always normal to the contour through that point. If selection acts directly on ϕ, then contours of equal phenotype are contours of equal fitness. Thus, the fitness gradient (∇w) points in the same direction as $\nabla\phi$.

The epistasis matrix in equation (2) captures the local nonlinearity of the phenotype function. Only the off-diagonal elements, which measure the degree to which changing one factor affects the phenotypic consequences of changing the other, correspond to developmental epistasis. The terms along the diagonal are more analogous to measures of dominance. We shall see, however, that the entire matrix emerges naturally from the study of evolution on a phenotype landscape. I will use systems with two underlying factors as examples; however, all results can be generalized to any number of dimensions (see Appendix). If there are n underlying factors, then E is an $n \times n$ matrix with elements

$$E_{ij} = \frac{\partial^2\phi}{\partial u_i \partial u_j}. \tag{3}$$

To find points of maximum canalization, we must find the minima of slope along a particular contour. The slope, $C(u_1, u_2)$, at a particular point is the magnitude of the phenotype gradient,

$$C_{u_1,u_2} = \|\nabla\phi\| = \sqrt{\left(\frac{\partial\phi}{\partial u_1}\right)^2 + \left(\frac{\partial\phi}{\partial u_2}\right)^2}. \tag{4}$$

We seek a maximum or minimum of C along a contour of constant ϕ. By the method of Lagrange multipliers, such a point will satisfy $\nabla C = k\nabla\phi$, where k is a number. ∇C is given by

$$\nabla C = \frac{1}{C}\begin{bmatrix} \frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1^2} + \frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1 \partial u_2} \\ \frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1 \partial u_2} + \frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_2^2} \end{bmatrix}, \tag{5}$$

so a point, (u_1', u_2'), is a maximum or minimum of canalization when

$$\left.\begin{aligned} \lambda\frac{\partial\phi}{\partial u_1} &= \frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1^2} + \frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1 u_2} \\ \lambda\frac{\partial\phi}{\partial u_2} &= \frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1 \partial u_2} + \frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_2^2} \end{aligned}\right|_{\substack{u_1=u_1' \\ u_2=u_2'}}, \tag{6}$$

where λ is a scalar that absorbs the $1/C$ term in the previous equation. (A good explication of the vector analysis used above and later in the paper is found in Marsden and Tromba 1996). We can write equation (6) in terms of the phenotype gradient and epistasis matrix:

$$\begin{bmatrix} \frac{\partial^2\phi}{\partial u_1^2} & \frac{\partial^2\phi}{\partial u_1 \partial u_2} \\ \frac{\partial^2\phi}{\partial u_1 \partial u_2} & \frac{\partial^2\phi}{\partial u_2^2} \end{bmatrix}\begin{bmatrix} \frac{\partial\phi}{\partial u_1} \\ \frac{\partial\phi}{\partial u_2} \end{bmatrix} = \lambda\begin{bmatrix} \frac{\partial\phi}{\partial u_1} \\ \frac{\partial\phi}{\partial u_2} \end{bmatrix}. \tag{7}$$

Thus, points of maximum or minimum canalization around a phenotype value occur precisely where the phenotype gradient is an eigenvector of the epistasis matrix. Figure 2 shows how these vectors relate to one another as we approach a canalized point.

Note that the vector on the right-hand side of equation (5) is the epistatis matrix, E, multiplied by the phenotype gradient. We thus write the general relation between the epistasis matrix, phenotype gradient, and gradient in slope (which points in the direction of decanalization) as

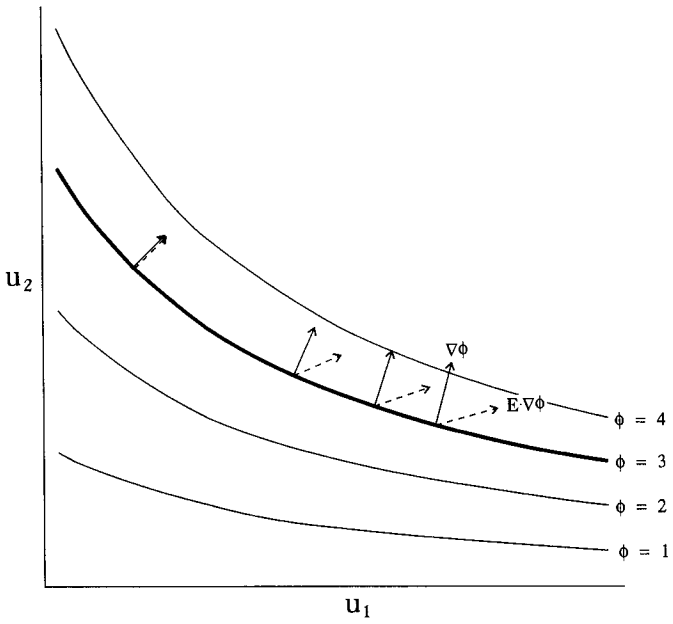

FIG. 2. Contours of the phenotype landscape for the function $\phi = u_1u_2$. This is a closeup view to illustrate the relation of $\nabla\phi$ (solid arrows) and $E{\cdot}\nabla\phi$ (dashed arrows). Note that as we approach a point of maximum canalization along a contour, these vectors become parallel.

$$E = \frac{\nabla\phi}{\|\nabla\phi\|} = \nabla C, \tag{8}$$

where $\|\nabla\phi\|$ is the length of $\nabla\phi$, it is always a positive number. ∇C is a vector pointing in the direction of maximum increase in slope, so we should expect selection for canalization to move the system in the opposite direction, along $-E{\cdot}\nabla\phi$. (See the Appendix for the n-dimensional derivation).

Before bringing selection into the picture, it is worth noting what equation (8) says about the relationship between epistasis and canalization. Recall that a contour on the phenotype landscape corresponds to a particular value of the phenotype. We can think of canalization around a particular phenotypic value as movement along a contour to a point of minimum slope (Fig. 1). $\nabla\phi$ is always normal to a contour, so increase in canalization is possible only when the vector $E{\cdot}\nabla\phi$ is not a scalar multiple of the vector $\nabla\phi$. In other words, when E is not a multiple of the identity matrix (a matrix with ones along the diagonal and zeros everywhere else). Note that this does not require that the off-diagonal elements of E be nonzero, only that if they are all zero, then the diagonal elements must not all be the same. Thus, epistasis is not necessary for canalization. It is, however, sufficient—because any nonzero off-diagonal elements preclude E from being the identity matrix (see Appendix).

Canalization is a property of the developmental system, not the kind of selection that the system is experiencing. To see when it will evolve, though, we need to bring selection into the discussion.

Evolution on a Phenotype Landscape

Define $w(\phi)$ as a fitness function (frequency and density independent) acting directly on the phenotype. If selection on the phenotype is not too strong, then we can approximate this function in the neighborhood of a point ϕ^* as

$$w(\phi^* + \delta_\phi) \approx w(\phi^*) + \delta_\phi \frac{dw}{d\phi} + \frac{\delta_\phi^2}{2}\frac{d^2w}{d\phi^2}. \tag{9}$$

In the following discussion, I will treat the underlying factors, u_i, as themselves being quantitative developmental characters (similar to Riska 1986 and Slatkin 1987) and assume that the phenotype landscape is approximately quadratic over the range of values of u_i present in the population. I further assume that the population is large enough that we may ignore drift and that the distribution of underlying factors in a population is multivariate normal. The assumption of normality is particularly important because it ensures that the distribution is symmetrical. I consider here only a single phenotypic character. The multicharacter case, in which each character has its own landscape, will be considered in another paper.

If we have a particular function in mind for $\phi(u_1, u_2)$, and there is no variation resulting from environmental effects, then we can capture the local geometry of the phenotype landscape by expanding in a Taylor series about $\phi(u_1, u_2)$. Letting

$$\delta_\phi = \phi_{u_1+x_1, u_2+x_2} - \phi_{u_1, u_2} \tag{10}$$

we expand around $\phi(u_1, u_2)$ and calculate δ_ϕ as

$$\delta_\phi(x_1, x_2) \approx x_1\frac{\partial\phi}{\partial u_1} + x_2\frac{\partial\phi}{\partial u_2} + \frac{x_1^2}{2}\frac{\partial^2\phi}{\partial u_1^2} + \frac{x_2^2}{2}\frac{\partial^2\phi}{\partial u_2^2} + x_1x_2\frac{\partial^2\phi}{\partial u_1\partial u_2}. \tag{11}$$

In any real population, different individuals will experience slightly different environmental conditions during development, adding an environmental component to variance in ϕ in addition to that resulting from variation in the u_i terms. In this case, the partial derivatives in equation (11) can be replaced by the coefficients of a quadratic least squares regression of ϕ on the u_i terms. This essentially finds a best-fit quadratic surface to the distribution of phenotypic values.

Now consider a population represented by a distribution of points in the space of underlying factors. We wish to find the change in the mean of this distribution over one generation with selection. Letting $\bar{u}_1$ and $\bar{u}_2$ be the mean values of u_1 and u_2, and $p(u_1, u_2)$ be the probability density associated with the distribution; then after selection, the change in the population mean in each of the u_i is given by

$$\Delta\bar{u}_i = h_i^2 \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\phi_{\bar{u}_1+x_1,\bar{u}_2+x_2})p_{\bar{u}_1+x_1,\bar{u}_2+x_2}x_1\,dx_1\,dx_2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\phi_{\bar{u}_1+x_1,\bar{u}_2+x_2})p_{\bar{u}_1+x_1,\bar{u}_2+x_2}\,dx_1\,dx_2}, \tag{12}$$

where h_i^2 denotes heritability of u_i. The denominator in equation (12) is simply the mean population fitness, $\bar{w}$.

Underlying factors that we measure are likely to be correlated with one another and show different variances. We can deal with this, with no loss of generality, by rotating the

axes so that the u_i are measured along the eigenvectors of the covariance matrix for the measured underlying characters and scaling these new axes to make variances equal. For purposes of iterating the evolution equations over multiple generations, I shall assume that any correlations between underlying factors are due to pleiotropy, rather than linkage, and that these change slowly.

Assuming that the probability density, $p(u_1, u_2)$, is bivariate normal, and writing the values $\bar{u}_1$ and $\bar{u}_2$ as a vector, $\bar{\mathbf{u}}$, integrating equation (12) yields

$$\Delta\bar{u} = \frac{H}{\bar{w}}\left[\sigma^2\frac{dw}{d\phi}\nabla\phi + \sigma^4\frac{d^2w}{d\phi^2}E\nabla\phi + \frac{\sigma^4}{2}\frac{d^2w}{d\phi^2}\nabla\phi\cdot\nabla^2\phi\right], \quad (13)$$

where σ^2 is the variance of the distribution of underlying factors (scaled to all be equal), $\mathbf{H}$ is a matrix with the h_i^2 along the diagonal and zeroes elsewhere, E is the epistasis matrix from equation (2), and $\nabla^2\phi$ is the Laplacian, defined as

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial u_1^2} + \frac{\partial^2\phi}{\partial u_2^2} \quad (14)$$

(see Appendix for derivation).

Note that the general evolution equation, without the assumption of normality involved in deriving equation (13) from equation (12), is obtained by substituting integrals (A5) and (A4) in the Appendix into equation (12) and interpreting the terms in the integrals as moments of the distribution of underlying factors.

By substituting equation (11) into equation (9), we are effectively mapping the values of the underlying factors to fitness. We could thus represent equations (12) and (13) as describing evolution on a (rather complicated) fitness landscape. Note that this fitness landscape would have a ridge running along the path followed by the optimum contour on the phenotype landscape. Dobzhansky (1937) suggested that adaptive landscapes may be characterized by ridges with local peaks along them, and the idea has recently gained renewed attention (Wagner et al. 1994; Gavrilets and Hastings 1995). We see here that this is the expected structure for adaptive landscapes associated with developmental characters.

The form of an adaptive landscape for developmental characters is determined by a combination of ecological factors (determining how selection acts on adult phenotype) and developmental interactions. The value of modeling evolution on a phenotype landscape is that it allows us to disentangle ecology from development.

Analysis and Discussion

Equation (13) describes how selection acting on a phenotypic character drives evolution of the underlying developmental or genetic factors that interact to construct that character. The result holds for any number of underlying factors. I have written this equation in an expanded form so that we can consider the various terms individually.

The first term on the right side of equation (13) is the standard result from evolutionary genetics; selection moves a population along the gradient in fitness ($dw/d\phi\cdot\nabla\phi$) at a rate proportional to the product of the slope of the fitness function and the amount of heritable variation within the population. If all effects are additive, then this term alone defines evolutionary dynamics. In this case the population takes the quickest route uphill or downhill to get to the optimum phenotype, the contour of which is a straight line (or, in n dimensions, a flat $n - 1$ dimensional surface) within which all points are neutrally stable.

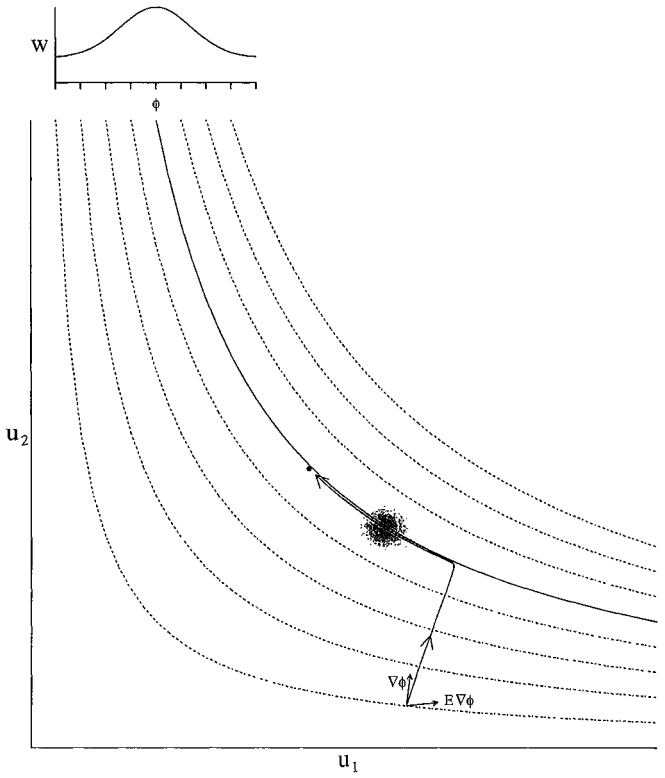

FIG. 3. An evolutionary trajectory on the same landscape as discussed in Figure 2 (here showing the entire landscape). The shaded area represents the population distribution during the stage in which it is undergoing canalization. Here, fitness drops off as a Gaussian function around the optimum phenotype (solid contour), such that $d^2w/d\phi^2$ is positive far from the optimum and negative near it. Note that at the outset, the trajectory veers away from the phenotype gradient in the direction of $E\cdot\nabla\phi$. Once near the optimum, the trajectory parallels it until reaching a local canalized point.

The second term captures selection for or against canalization. Recall that the vector $E\cdot\nabla\phi$ points in the direction of maximum increase in the slope of the phenotype function (Fig. 2). When the system is near an optimum phenotype, such that $d^2w/d\phi^2$ is negative (as it must be at a stable selective equilibrium), then this term represents selection for canalization, shifting the population toward a local minimum in the slope of the phenotype landscape. By contrast, if $d^2w/d\phi^2$ is positive, then the trajectory is biased in the direction of $E\cdot\nabla\phi$. In other words, the population moves away from a canalized state toward regions of maximum sensitivity of phenotype to variation in the underlying factors. These different phases in the response of development to selection are illustrated in Figures 3 and 4.

Lande and Arnold (1983) obtained a similar result, defining $d^2w/d\phi^2$ as the strength of stabilizing (or destabilizing) selection. We see here that how a population responds to such

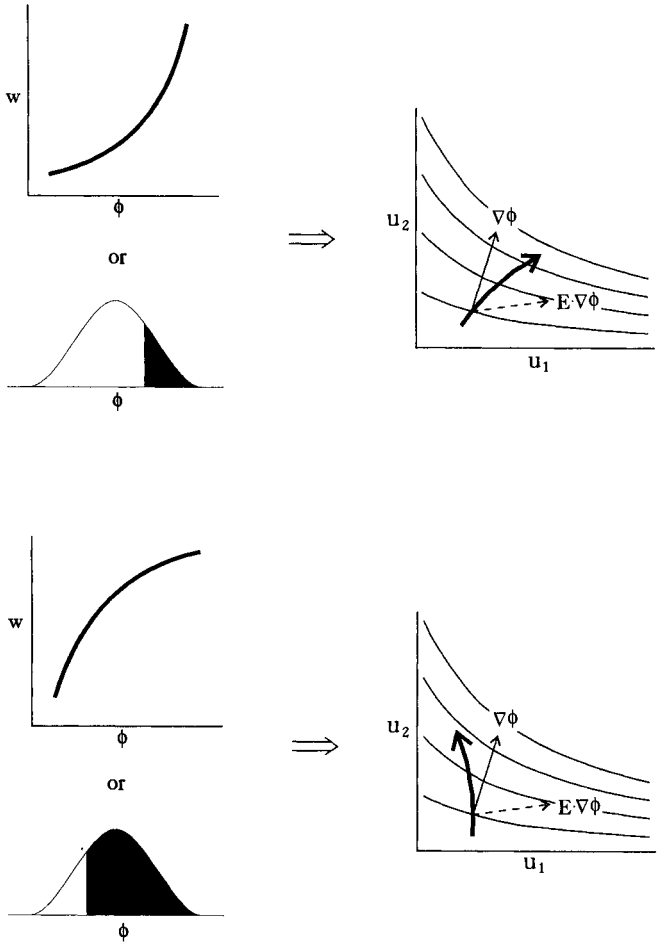

FIG. 4. How the selection regime influences the dynamics of evolution. The upper figures on the left represent a continuous fitness function acting on the phenotype. Below them are figures representing truncation selection in which the darkened region of the distribution is selected to produce the next generation. An accelerating fitness function (top) or strong truncation selection shifts the trajectory toward steeper parts of the phenotype landscape. Decelerating selection has the opposite effect.

selection depends on the process by which ϕ develops. If the vector $E\cdot\nabla\phi$ has a component pointing along a phenotype contour, then stabilizing selection leads to canalization.

The third term in equation (13) shows that the equilibrium state of the system can shift away from the optimum phenotype if the phenotype landscape curves away more abruptly on one side of the optimum than on the other. This curvature of the landscape induces a skewed distribution on ϕ. Because $\nabla^2\phi$ is a number, this last term is a vector pointing along the phenotype gradient. It thus does not contribute to canalization.

Figure 3 shows a trajectory resulting from iteration of equation (13) for the phenotype landscape defined by $\phi = u_1u_2$ (the simplest landscape that allows canalization, used here for illustrative purposes only). In this and the following examples, I am holding the value of σ^2 constant as the system evolves. This term is likely to change in accordance with the strength and kind of selection acting. Holding it constant over short periods, though, is consistent with the expectation that the variance of a distribution of allelic effects changes much more slowly than the mean (Lande 1976) and allows us to focus on the effects of developmental interactions on evolutionary dynamics.

Note that for the example shown in Figure 3, if the goal were simply to model the change in the mean value of the character, ϕ, a model that considers only additive effects (i.e., that follows the gradient) would do quite well. If we wish to study the evolution of development, though, such a model would be insufficient. As the figure shows, the initial approach to the optimum phenotype contour is followed by a comparable amount of developmental change as the character becomes canalized. This second phase of evolution results from nonlinear interactions between the underlying factors, manifest as a change in slope along the contour.

Our principal interest is in the impact of phenotypic selection on developmental/genetic parameters. For this, we return to the second term in equation (13) and the fact that epistasis can cause the trajectory of evolution to veer away from the gradient of fitness. Figure 4 shows the relation between the shape of the fitness function and its effect on developmental evolution. A decelerating fitness function causes the trajectory to shift toward the less-steep region of the surface, corresponding to canalizing selection on development. Conversely, accelerating selection on phenotype produces decanalizing selection. (What I am calling decelerating and accelerating selection have been called, respectively, concave and convex by Templeton 1981.)

We can understand this intuitively by noting that with an accelerating fitness function, mean population fitness increases with increasing variance (holding the mean constant), because the benefit of deviating from the mean in the adaptive direction is greater than the cost of deviating in the other direction (Layzer 1980). A population subjected to such a fitness function thus tends to shift toward steeper parts of the phenotype landscape.

We generally assume that fitness functions are decelerating near an optimum phenotype (if the function is continuous, this is a requirement for an optimum), and Gillespie (1978) and Templeton (1981) have argued that fitness functions are generally concave (decelerating). These arguments were based on assumptions about enzyme kinetics, and may not apply to selection on general phenotypic characters. It seems likely that selection is canalizing most of the time. Occasional decanalizing selection, though, can have significant consequences for developmental evolution. The degree to which selection is canalizing or decanalizing is independent of absolute fitness. Thus, a population could experience decanalizing selection without being in danger of extinction.

To understand when we might expect divergence in development under uniform selection, consider a population that is at a canalized point for some optimum phenotype (Fig. 5). If the optimum changes, so that the population now experiences accelerating selection, what was a stable canalized point becomes an unstable bifurcation point, because movement in any direction along a contour increases slope. Nearby populations can thus diverge from one another in the space of underlying developmental factors, even as they show parallel evolution of the character under selection. This effect

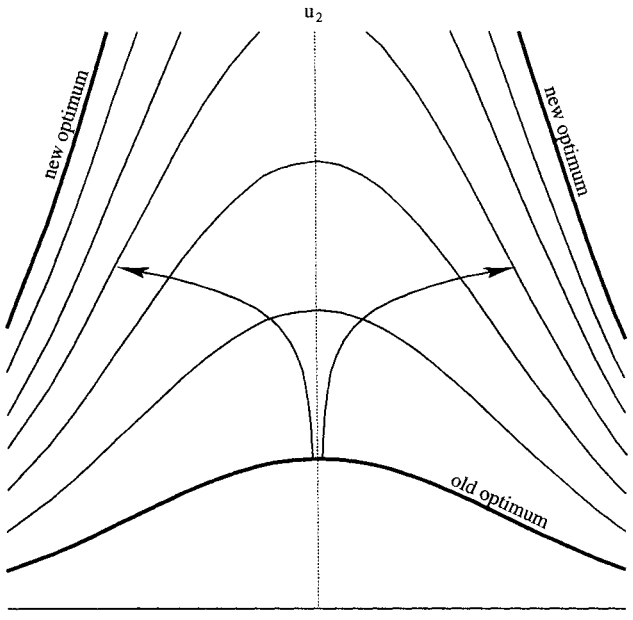

FIG. 5. Evolution away from a canalized point on the phenotype landscape defined by $\phi = u_2 exp(u_1^3)$. The combined effects of epistasis and a diverging gradient field cause nearby trajectories to diverge markedly. Note that the trajectories are definitely veering away from the gradient, which is perpendicular to the contours of equal phenotype. The fitness function here is Gaussian.

is compounded when the gradient field ($\nabla\phi$) itself diverges in the direction in which the population is evolving. This is the case in Figure 5; the two populations would diverge somewhat, though not as much, even if they followed the gradient of the phenotype surface. We thus arrive at the surprising result that stabilizing selection moves the system to a point that is poised to become maximally unstable should it be subjected to decanalizing selection.

In addition to showing us how to break von Baer's laws, this result has implications for speciation. Two isolated populations may respond very differently at the genetic or developmental level to the same selection regime. They thus may land at very different points on the new optimum phenotype contour. If there are multiple points of maximum canalization along that contour, subsequent stabilizing selection could drive them apart (Barton 1996), and different ways of achieving that phenotype may not be compatible if they are later crossed, leading to reduced fitness among hybrids (Fig. 6).

Wright (1977) and Layzer (1980) also noted that a large shift in the optimum phenotype can cause selection for increased variance, which could lead to divergence. We see here that the potential for such selection to drive initially similar populations apart is greatest when the populations are initially highly canalized. This suggests an empirical test for this particular hypothesis for divergence: the degree to which isolated parts of a population diverge under accelerating selection should be reduced if the initial population from which they are drawn is first moved away from a canalized state, perhaps by being subjected to strong directional selection or crossed with a different population.

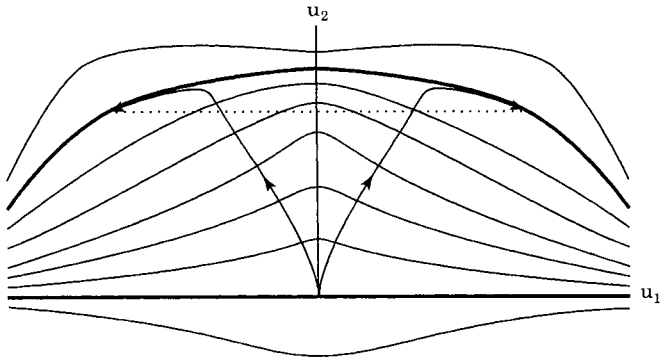

FIG. 6. Decanalizing selection can cause initially similar populations to diverge permanently if the new optimum phenotype has multiple canalized points.

The results presented here sit well with the observation that von Baer's laws are usually true, but are sometimes broken spectacularly. There is one environment, though, where decanalizing selection is the norm: the laboratory. Truncation selection can have the same decanalizing effect as a continuous fitness function that is accelerating (Fig. 4). This occurs when the selection threshold is set so that less than 50% of the population is chosen to produce the next generation (see Appendix). This is a common strategy in laboratory selection experiments, many of which are thus imposing decanalizing selection while studying the evolution of some particular character.

This suggests another empirical test. If decanalizing selection is in fact responsible for some of the genetic divergence under uniform selection seen in experiments such as those of Cohan (1984) and Cohan and Hoffmann (1989), then the theory presented here predicts that such divergence could be greatly reduced simply by resetting the selection threshold.

Figure 5 illustrates one possible mechanism for the breaking of von Baer's laws. There are certainly others. For instance, if selection for canalization is relatively weak compared to stabilizing selection on phenotype, then genetic drift can move a population along a contour of equal phenotype. Equation (13) shows that whereas the strength of directional selection goes as σ^2, the population variance, the strength of canalizing selection goes as σ^4. This mirrors the results of Wagner et al. (1997), who found that, in simulations, the degree of canalization that arose increased abruptly with increasing population variance. Selection holding a population at a point of maximum canalization along a phenotype contour is thus much more sensitive to population variance than is the selection moving a population toward that contour.

We can approximate evolution along an optimum phenotype contour by the component of $E\cdot\nabla\phi$ that is parallel to that contour. This term, Z, is the component of canalizing selection that is independent of directional selection. It is given by

$$Z = \frac{\langle E\nabla\phi, \nabla\phi\rangle}{\|\nabla\phi\|^2}\nabla\phi - E\nabla\phi \qquad (15)$$

(see Appendix for derivation). Here, $\langle E\cdot\nabla\phi, \nabla\phi\rangle$ is the inner product between the two vectors, which is related to the angle

between them—being zero when the vectors are perpendicular to one another.

Note that if E is a multiple of the identity matrix, then selection will not move a population along a phenotype contour. A population on such a landscape with n underlying factors is thus free to drift along a contour of dimension $n - 1$. This was noted by Lande (1976) for a phenotype resulting from additive allelic effects ($E = 0$). We see here that the same thing happens with nonadditive effects as long as the diagonal elements of E are all equal and the off-diagonal elements are all zero. The contours on such a surface are concentric circles (or spheres if $n > 3$), so the maximum slope ($\|\nabla\phi\|$) is the same at each point around a contour.

The above examples assume that the shape of the distribution of underlying factors does not change over a number of generations. While this may hold approximately for weak selection (Lande 1976), we might expect that the moments of the distribution will evolve in response to stronger selection (Wagner 1989; Turelli and Barton 1994). There are two consequences of this. First, selection could impose covariation or unequal variances of the underlying factors even if the distribution remains multivariate normal. Second, and more importantly, the distribution could evolve to be asymmetrical, with a nonzero third moment.

Changing the variances of the us or inducing covariation can alter the rate at which canalization evolves and require that we again transform the axes before applying equation (13) for the next generation. This will not change the important aspects of the landscape, though, because the relationships between E and $\nabla\phi$ are invariant under rotation and points of maximum canalization cannot be eliminated by uniformly stretching the axes.

Skewing the distribution of underlying factors has a different effect. By altering the way that the population samples the phenotype landscape, a skewed distribution of the underlying factors could shift the equilibrium slightly away from a point of maximum canalization.

Note that this model does allow the distribution of ϕ to be skewed and this can result in the population mean being shifted slightly off of an optimum contour. In fact, both the second and third terms on the right side of equation (13) may be thought of as describing selection to change the form of the phenotype distribution. The second term captures selection to decrease the variance in ϕ (canalization) or to increase it (decanalization). The third term captures selection to skew the distribution of ϕ. Because I am constraining the distribution of underlying factors, the only way to achieve a change in the phenotype distribution is to move on the phenotype landscape.

Since the underlying factors and the character that they build are all quantitative characters, we could model their joint evolution using a variance/covariance matrix (Cheverud 1984). Treating the partial derivatives in equation (11) as regression coefficients, equation (13) does essentially the same thing, but replaces the covariance of ϕ and u_i with a function derived from a hypothesized causal relationship in development. The phenotype landscape approach thus makes use of the fact that we know *why* ϕ and u_i covary. This allows us to describe phenotypic evolution in terms of the developmental changes that must accompany it and to see how developmental processes can evolve even when the average phenotype that they construct does not change.

Acknowledgments

I thank G. P. Wagner, J. Kim, L. Buss, and J. M. Cheverud for helpful discussion of the ideas and the manuscript.

Literature Cited

Barton, N. H. 1996. Natural selection and random genetic drift as causes of evolution on islands. Phil. Trans. R. Soc. Lond. B. 351:785–795.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol. 110:155–171.

Cheverud, J. M., and E. J. Routman. 1996. Epistasis as a source of increased additive genetic variance at population bottlenecks. Evolution 50:1042–1051.

Cohan, F. M. 1984. Genetic divergence under uniform selection. I. Similarity among populations of *Drosophila melanogaster* in their response to artificial selection for modifiers of ci^D. Evolution 38:55–71.

Cohan, F. M., and A. A. Hoffmann. 1989. Uniform selection as a diversifying force in evolution: evidence from *Drosophila*. Am. Nat. 134:613–637.

Darwin, C. 1859. On the origin of species by means of natural selection. Murray, London.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

Freeman, G., and J. W. Lundelius. 1992. Evolutionary implications of the mode of D quadrant specification in coelomates with spiral cleavage. J. Evol. Biol. 5:205–247.

Gavrilets, S., and A. Hastings. 1995. Founder effect speciation: a theoretical reassessment. Am. Nat. 147:466–491.

Gilbert, S. F. 1991. A conceptual history of modern embryology. Johns Hopkins Univ. Press, Baltimore, MD.

Gillespie, J. H. 1978. A general model to account for enzyme variation in natural populations. II. Characterization of the fitness functions. Am. Nat. 110:809–821.

Hall, B. K. 1992. Evolutionary developmental biology. Chapman & Hall, New York.

Jacobson, A. G., and A. K. Sater. 1988. Features of embryonic development. Development 104:341–359.

Kluge, A. G., and R. E. Strauss. 1985. Ontogeny and systematics. Ann. Rev. Ecol. Syst. 16:247–268.

Lande, R. 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. 26:221–235.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Layzer, D. 1980. Genetic variation and progressive evolution. Am. Nat. 115:809–826.

Levinton, J. S. 1988. Genetics, paleontology, and macroevolution. Cambridge Univ. Press, Cambridge, U.K.

Marsden, J. E., and A. J. Tromba. 1996. Vector calculus. 4th ed. Freeman, New York.

Raff, R. A. 1996. The shape of life. Univ. of Chicago Press, Chicago.

Riska, B. 1986. Some models for development, growth, and morphometric correlation. Evolution 40:1303–1311.

Rutledge, J. J., E. J. Eisen, and J. E. Legates. 1974. Correlated response in skeletal traits and replicate variation in selected lines of mice. Theor. Appl. Genet. 45:26–31.

Simpson, G. G. 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.

Slatkin, M. 1987. Quantitative genetics of heterochrony. Evolution 41:799–811.

Sterns, S. C., and T. J. Kawecki. 1994. Fitness sensitivity and the canalization of life history traits. Evolution 48:1438–1450.

Templeton, A. R. 1981. Some comments on "genetic variation and progressive evolution" by D. Layzer. Am. Nat. 117:1049–1051.

TURELLI, M., AND N. H. BARTON. 1994. Genetic and statistical analysis of strong selection on polygenic traits: What, me normal? Genetics 138:913–941.
VON BAER, K. E. 1828. Über entwicklungsphysiologische der thiere, beobachtung und reflexion. Gebrüder Bornträger, Königsberg, Germany.
WADDINGTON, C. H. 1957. The strategy of the genes. MacMillan, New York.
WAGNER, A., G. P. WAGNER, AND P. SIMILION. 1994. Epistasis can facilitate the evolution of reproductive isolation by peak shifts: a two-locus two-allele model. Genetics 138:533–545.
WAGNER, G. P. 1989. Multivariate mutation-selection balance with constrained pleiotropic effects. Genetics 122:223–234.
WAGNER, G. P., G. BOOTH, AND H. BAGHERI-CHAICHIAN. 1997. A population genetic theory of canalization. Evolution 51:329–347.
WRAY, G. A. 1994. The evolution of cell lineage in echinoderms. Am. Zool. 34:353–363.
WRAY, G. A., AND R. A. RAFF. 1990. Novel origins of lineage founder cells in the direct-developing sea urchin *Heliocidaris erythrogramma*. Dev. Biol. 141:41–54.
WRIGHT, S. 1977. Evolution and the genetics of populations. Vol. 3. Experimental results and evolutionary deductions. Univ. of Chicago Press, Chicago.

Corresponding Editor: L. Stevens

APPENDIX

n-Dimensional Landscapes

For the general case of n underlying factors contributing to a character

$$C = \|\nabla\phi\| = \sqrt{\sum_{i=0}^{n-1}\left(\frac{\partial\phi}{\partial u_i}\right)^2} \tag{A1}$$

so,

$$\nabla C = \begin{bmatrix} \frac{\partial C}{\partial u_1} \\ \frac{\partial C}{\partial u_2} \\ \vdots \\ \frac{\partial C}{\partial u_n} \end{bmatrix} = \frac{1}{C}\begin{bmatrix} \sum_i \frac{\partial\phi}{\partial u_i}\frac{\partial^2\phi}{\partial u_1\partial u_i} \\ \sum_i \frac{\partial\phi}{\partial u_i}\frac{\partial^2\phi}{\partial u_2\partial u_i} \\ \vdots \\ \sum_i \frac{\partial\phi}{\partial u_i}\frac{\partial^2\phi}{\partial u_n\partial u_i} \end{bmatrix}. \tag{A2}$$

The vector of sums on the far right of equation (A2) is simply $E\cdot\nabla\phi$, so we get equation (8) again.

To show that epistasis is sufficient for canalization to be possible, we use the fact that E is symmetric. A result of this is that E has distinct, orthogonal eigenvectors that we can arrange into a new matrix, S, such that $S^{-1}ES = \Lambda$, where Λ is a matrix with the eigenvectors of E along the diagonal and zeros everywhere else. Canalization is possible so long as E cannot be diagonalized to produce the identity matrix. In other words, so long as $\Lambda \neq I$. Suppose that $\Lambda = I$, then

$$\begin{aligned} S^{-1}ES &= I \\ E &= SIS^{-1} \\ E &= SS^{-1} \\ E &= I. \end{aligned} \tag{A3}$$

Therefore, E is the identity matrix if $\Lambda = I$, so any epistasis is sufficient to allow for canalization somewhere along a contour.

Derivation of the Evolution Equation

To derive equation (13), we substitute equation (11) into equation (9), then integrate as in equation (12). I shall consider the two-dimensional case first. For u_1, the upper integral in equation (12) is

APPENDIX. Continued.

$$\begin{aligned} \int\int \Bigg\{ & w(\phi)x_1 + \frac{dw}{d\phi}\Bigg(x_1^2\frac{\partial\phi}{\partial u_1} + x_1x_2\frac{\partial\phi}{\partial u_2} + \frac{x_1^3}{2}\frac{\partial^2\phi}{\partial u_1^2} + \frac{x_1x_2^2}{2}\frac{\partial^2\phi}{\partial u_2^2} \\ & + x_1^2x_2\frac{\partial^2\phi}{\partial u_1\partial u_2}\Bigg) \\ & + \frac{1}{2}\frac{d^2w}{d\phi^2}\Bigg[x_1^3\left(\frac{\partial\phi}{\partial u_1}\right)^2 + x_1x_2^2\left(\frac{\partial\phi}{\partial u_2}\right)^2 + \frac{x_1^5}{4}\left(\frac{\partial^2\phi}{\partial u_1^2}\right)^2 \\ & + \frac{x_1x_2^4}{4}\left(\frac{\partial^2\phi}{\partial u_2^2}\right)^2 + x_1^3x_2^2\left(\frac{\partial^2\phi}{\partial u_1\partial u_2}\right)^2 + 2x_1^2x_2\frac{\partial\phi}{\partial u_1}\frac{\partial\phi}{\partial u_2} \\ & + x_1^4\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1^2} + x_1^2x_2^2\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_2^2} + 2x_1^3x_2\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1\partial u_2} \\ & + x_1^3x_2\frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1^2} + x_1x_2^3\frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_2^2} + 2x_1^2x_2^2\frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1\partial u_2} \\ & + \frac{x_1^3x_2^2}{2}\frac{\partial^2\phi}{\partial u_1^2}\frac{\partial^2\phi}{\partial u_2^2} + x_1^4x_2\frac{\partial^2\phi}{\partial u_1^2}\frac{\partial^2\phi}{\partial u_1\partial u_2} \\ & + x_1^2x_2^3\frac{\partial^2\phi}{\partial u_2^2}\frac{\partial^2\phi}{\partial u_1\partial u_2}\Bigg]\Bigg\}p(x_1,x_2)\,dx_1\,dx_2, \end{aligned} \tag{A4}$$

and for u_2,

$$\begin{aligned} \int\int \Bigg\{ & w(\phi)x_2 + \frac{dw}{d\phi}\Bigg(x_1x_2\frac{\partial\phi}{\partial u_1} + x_2^2\frac{\partial\phi}{\partial u_2} + \frac{x_1^2x_2}{2}\frac{\partial^2\phi}{\partial u_1^2} + \frac{x_2^3}{2}\frac{\partial^2\phi}{\partial u_2^2} \\ & + x_1x_2^2\frac{\partial^2\phi}{\partial u_1\partial u_2}\Bigg) \\ & + \frac{1}{2}\frac{d^2w}{d\phi^2}\Bigg[x_1^2x_2\left(\frac{\partial\phi}{\partial u_1}\right)^2 + x_2^3\left(\frac{\partial\phi}{\partial u_2}\right)^2 + \frac{x_1^4x_2}{4}\left(\frac{\partial^2\phi}{\partial u_1^2}\right)^2 \\ & + \frac{x_2^5}{4}\left(\frac{\partial^2\phi}{\partial u_2^2}\right)^2 + x_1^2x_2^3\left(\frac{\partial^2\phi}{\partial u_1\partial u_2}\right)^2 + 2x_1x_2^2\frac{\partial\phi}{\partial u_1}\frac{\partial\phi}{\partial u_2} \\ & + x_1^3x_2\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1^2} + x_1x_2^3\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_2^2} + 2x_1^2x_2^2\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1\partial u_2} \\ & + x_1^2x_2^2\frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1^2} + x_2^4\frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_2^2} + 2x_1x_2^3\frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1\partial u_2} \\ & + \frac{x_1^2x_2^3}{2}\frac{\partial^2\phi}{\partial u_1^2}\frac{\partial^2\phi}{\partial u_2^2} + x_1^3x_2^2\frac{\partial^2\phi}{\partial u_1^2}\frac{\partial^2\phi}{\partial u_1\partial u_2} \\ & + x_1x_2^4\frac{\partial^2\phi}{\partial u_2^2}\frac{\partial^2\phi}{\partial u_1\partial u_2}\Bigg]\Bigg\}p(x_1,x_2)\,dx_1\,dx_2. \end{aligned} \tag{A5}$$

The key to integrating the result is to use the facts that, under the assumptions in the text, only the terms in x_i^2, x_i^4, and $x_i^2x_j^2$ are nonzero. These terms correspond to the moments of the normal distribution.

$$\begin{aligned} \int\int x_i^2p(x_i,x_j)\,dx_i\,dx_j &= \sigma^2 \\ \int\int x_i^2x_j^2p(x_i,x_j)\,dx_i\,dx_j &= \sigma^4 \\ \int\int x_i^4p(x_i,x_j)\,dx_i\,dx_j &= 3\sigma^4. \end{aligned} \tag{A6}$$

Recall that we are assuming equal variance in all characters and no correlations. Integrating yields

APPENDIX. Continued.

$$\Delta\overline{u_1} = \sigma^2\frac{dw}{d\phi}\frac{\partial\phi}{\partial u_1} + \sigma^4\frac{d^2w}{d\phi^2}\left(\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1^2} + \frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_1\partial u_2}\right) + \frac{\sigma^4}{2}\frac{d^2w}{d\phi^2}\frac{\partial\phi}{\partial u_1}\left(\frac{\partial^2\phi}{\partial u_1^2} + \frac{\partial^2\phi}{\partial u_2^2}\right) \quad \text{(A7)}$$

and

$$\Delta\overline{u_2} = \sigma^2\frac{dw}{d\phi}\frac{\partial\phi}{\partial u_2} + \sigma^4\frac{d^2w}{d\phi^2}\left(\frac{\partial\phi}{\partial u_1}\frac{\partial^2\phi}{\partial u_1\partial u_2} + \frac{\partial\phi}{\partial u_2}\frac{\partial^2\phi}{\partial u_2^2}\right) + \frac{\sigma^4}{2}\frac{d^2w}{d\phi^2}\frac{\partial\phi}{\partial u_2}\left(\frac{\partial^2\phi}{\partial u_1^2} + \frac{\partial^2\phi}{\partial u_2^2}\right). \quad \text{(A8)}$$

Rearranging these gives equation (13).

For the n-dimensional case:

$$\delta_\phi = \sum_i x_i\frac{\partial\phi}{\partial u_i} + \frac{1}{2}\sum_i\sum_j x_ix_j\frac{\partial^2\phi}{\partial u_i\partial u_j} \quad \text{(A9)}$$

and

$$\delta_\phi^2 = \sum_i\sum_j x_ix_j\frac{\partial\phi}{\partial u_i}\frac{\partial\phi}{\partial u_j} + \frac{1}{4}\sum_i\sum_j\sum_k\sum_l x_ix_jx_kx_l\frac{\partial^2\phi}{\partial u_i\partial u_j}\frac{\partial^2\phi}{\partial u_k\partial u_l} + \sum_i\sum_j\sum_k x_ix_jx_k\frac{\partial\phi}{\partial u_i}\frac{\partial^2\phi}{\partial u_j\partial u_k}. \quad \text{(A10)}$$

For the ath factor, multiply by x_a and consider only terms that do not integrate to zero. This yields

$$x_a^2\frac{\partial\phi}{\partial u_a} + x_a^4\frac{\partial\phi}{\partial u_a}\frac{\partial^2\phi}{\partial u_a^2} + \sum_{i\neq a}x_a^2x_i^2\frac{\partial\phi}{\partial u_a}\frac{\partial^2\phi}{\partial u_i^2} + 2\sum_{i\neq a}x_a^2x_i^2\frac{\partial\phi}{\partial u_i}\frac{\partial^2\phi}{\partial u_a\partial u_i}. \quad \text{(A11)}$$

Integrating this, for each factor, yields equation (13).

Truncation Selection

To specify the conditions under which truncation selection will be decanalizing, consider a character, u, that is normally distributed with mean μ and variance σ^2. If we select every individual with

APPENDIX. Continued.

character u greater than some threshold value, s, then the proportion of the population selected is given by

$$w = \int_s^\infty \frac{1}{\sqrt{2\pi}\sigma}e^{[(u-\mu)^2]/2\sigma^2}\,du. \quad \text{(A12)}$$

If selection is decanalizing, then $dw/d\sigma > 0$, increasing the variance will increase the proportion of the population selected, so the part of the population that lies on a steeper part of the phenotype landscape will have higher fitness than the part lying on a less steep part.

Substituting $v = (u - \mu)/\sigma$ into equation (A12) gives

$$\int_{(s-\mu)/\sigma}^\infty \frac{1}{\sqrt{2\pi}}e^{v^2/2}\,dv. \quad \text{(A13)}$$

Let $\Phi(v)$ be the cumulative standard normal density function. Recalling that $\phi(\infty) = 1$, equation (A13) can be rewritten:

$$w = 1 - \Phi\left(\frac{s-\mu}{\sigma}\right). \quad \text{(A14)}$$

Setting $y = (s - \mu)/\sigma$ and differentiating with respect to σ yields

$$\frac{dw}{d\sigma} = -\frac{d\Phi(y)}{dy}\frac{dy}{d\sigma} = \frac{d\Phi(y)}{dy}\frac{s-\mu}{\sigma^2}. \quad \text{(A15)}$$

Since $\Phi(y)$ is a continuously increasing function of y, we find that

$$\frac{dw}{d\sigma} > 0 \quad \text{iff } s > \mu. \quad \text{(A16)}$$

Thus, such selection is decanalizing when we select less than the top 50% of the population.

Evolution along a Contour

To find the vector $\mathbf{Z}$, which is the component of $E\cdot\nabla\phi$ that lies along a phenotype contour, we use the fact that the projection of $E\cdot\nabla\phi$ onto a line in the direction of $\nabla\phi$ is a vector, p, given by

$$p = \frac{\langle E\nabla\phi, \nabla\phi\rangle}{\|\nabla\phi\|^2}\nabla\phi, \quad \text{(A17)}$$

where $\langle E\ \nabla\phi, \nabla\phi\rangle$ is the inner product of the two vectors. The direction of maximum increase in canalization along the phenotype contour is then given by a vector, $\mathbf{Z}$, pointing from $E\cdot\nabla\phi$ to p.