

# Developmental Associations Between Traits: Covariance and Beyond

Sean H. Rice<sup>1</sup>

*Department of Ecology and Evolutionary Biology, Osborn Memorial Labs, Yale University, New Haven, Connecticut 06520*

Manuscript received December 17, 2002  
Accepted for publication September 11, 2003

## ABSTRACT

Statistical associations between phenotypic traits often result from shared developmental processes and include both covariation between the trait values and more complex associations between higher moments of the joint distribution of traits. In this article, an analytical technique for calculating the covariance between traits is presented on the basis of (1) the distribution of underlying genetic and environmental variation that jointly influences the traits and (2) the mechanics of how these underlying factors influence the development of each trait. It is shown that epistasis can produce patterns of covariation between traits that are not seen in additive models. Applying this approach to a trait in parents and the same trait in their offspring allows us to study the consequences of epistasis for the evolution of additive genetic variance and heritability. This analysis is then extended to the study of more complicated associations between traits. It is shown that even traits that are not correlated may exhibit developmental associations that influence their joint evolution.

ONE of the most important ways in which development influences evolution is by producing statistical associations between different phenotypic traits that influence the joint evolution of those traits (MAYNARD SMITH *et al.* 1985). Most theoretical models of developmental associations between traits have focused on genetic covariance, resulting either from additive pleiotropic effects (LANDE 1979; ROFF 1997) or from particular models of development, usually involving differential allocation of resources to growing structures (RENDEL 1963; RISKÄ 1986; HOULE 1991). These models generally assume that any developmental association between two traits can be satisfactorily captured by the genetic covariance between them.

In this article, I present a way to model developmental associations between traits that is amenable to any sort of epistatic interactions between genetic and environmental factors influencing those traits. I apply this approach to the question of how epistatic interactions influence phenotypic and genetic covariation and to the study of developmental associations that are more complicated than covariance.

The analysis that follows is phrased in terms of phenotype landscapes (RICE 2002). A phenotype landscape is a map of some phenotypic trait as a function of all of the underlying genetic and environmental factors that contribute to it. In this theory, environmental factors (such as temperature at a particular stage of development, salinity, etc.) are treated as underlying factors that may not be heritable (although they can be). We thus do not

focus on the traditional decomposition of variation into additive genetic, dominance, and environmental components [although such a decomposition can be recovered from the phenotype landscape (WOLF *et al.* 2001) given a model of transmission]. In this article I assume that phenotype landscapes are continuous and infinitely differentiable. This assumption is for illustrative purposes only; a slight modification of the theory presented here applies as well to discontinuous landscapes (RICE 2002). A note on terminology: The term “epistasis” is used in a number of different ways in the literature (PHILLIPS 1998). Throughout this article, I use the word epistasis to refer to nonadditive interactions between underlying factors in their contributions to phenotype. Thus, the phenotype function  $\phi = au_1 + bu_2$ , where  $\phi$  is the value of a trait,  $u_1$  and  $u_2$  are values of genetic underlying factors, and  $a$  and  $b$  are constants, exhibits no epistasis since the underlying factors contribute additively to the trait value. By contrast, phenotype functions such as the following do exhibit epistasis:  $\phi = u_1u_2$ ,  $\phi = u_1/u_2$ ,  $\phi = u_1^2$ , etc.

The mathematical tools for the representation and analysis of phenotype landscapes are the same as those presented in RICE (2002). The local geometry of the phenotype landscape and the distribution of underlying genetic or environmental variation are captured with tensors. For our purposes, a tensor is an array of values (such as a vector or a matrix) that has the property that the relationship to other such arrays is unchanged if we rotate the coordinate axes (I discuss the biological meaning of this below). The rank of a tensor is the number of subscripts necessary to identify each element. Thus, a vector is a tensor of rank one and the matrices used in this article are tensors of rank two (nonsquare matrices are not tensors, but we have no use for these

<sup>1</sup>Address for correspondence: Department of Ecology and Evolutionary Biology, Osborn Memorial Labs, 165 Prospect St., Yale University, New Haven, CT 06520. E-mail: sean.rice@yale.edu

in this article). (The dimension of a tensor is just the number of values that each subscript can take; so a  $3 \times 3$  matrix has rank two, because two subscripts are required to identify any particular element, but dimension three, because each subscript can take any of three values).

The notation follows that in RICE (2002). Let  $\mathbf{P}^n$  be an  $n$ th rank tensor with elements being the  $n$ th moments of the distribution of underlying factors (throughout this article, “moments” refers to central moments). Defining  $x_i = (u_i - \bar{u}_i)$ , *i.e.*, each value of the  $i$ th trait measured relative to the population mean for that trait, the elements of  $\mathbf{P}^n$  are

$$\mathbf{P}^n(i_1, i_2, \dots, i_n) = E[x_{i_1} x_{i_2} \dots x_{i_n}]. \quad (1)$$

Because  $x_i = (\mathbf{u}_i - \bar{\mathbf{u}}_i)$ , the vector (or tensor of rank 1)  $\mathbf{P}^1$  is zero.  $\mathbf{P}^2$  is the standard covariance matrix, with elements  $\mathbf{P}_{ij}^2 = E(x_i x_j) = \text{Cov}(\mathbf{u}_i, \mathbf{u}_j)$ .  $\mathbf{P}^3$  is a third-order array of third moments. Higher-ranked  $\mathbf{P}$  tensors contain the higher-order moments; these include univariate terms such as kurtosis ( $\mathbf{P}_{1111}^4 = E[x_1^4]$ ) as well as higher-order mixed moments, which do not generally have names. An important property of the  $\mathbf{P}$  tensors is that they are highly symmetrical. For the familiar case of  $\mathbf{P}^2$  (the covariance matrix)  $\mathbf{P}_{12}^2 = E(x_1 x_2) = E(x_2 x_1) = \mathbf{P}_{21}^2$ , this represents just the fact that covariances are symmetrical. Similarly  $\mathbf{P}_{1222}^4 = \mathbf{P}_{2122}^4 = \mathbf{P}_{2212}^4 = \mathbf{P}_{2221}^4 = E(x_1 x_2^3)$ . In other words, it is the number of each kind of subscript that matters, not their order. Since the rank of a tensor is the same as the number of subscripts needed to identify any particular element, I sometimes drop the superscript indicating rank when this is obvious from the number of subscripts; thus  $\mathbf{P}_{ijkm} = \mathbf{P}_{ijkm}^4$ .

Let  $\mathbf{D}_a^n$  be a tensor of rank  $n$ , associated with phenotypic trait  $\phi_a$ , with elements defined as

$$\mathbf{D}_a^n(i_1, i_2, \dots, i_n) = \frac{\partial^n \phi_a}{\partial u_{i_1} \partial u_{i_2} \dots \partial u_{i_n}}. \quad (2)$$

In words,  $\mathbf{D}_a^n$  contains all of the  $n$ th derivatives of phenotypic trait  $\phi_a$  with respect to the underlying factors.  $\mathbf{D}^1$  is just the gradient vector,  $\nabla \phi$ .  $\mathbf{D}^2$  is a matrix of second partial derivatives of phenotype with respect to the underlying factors [this is the same as the matrix  $E$  in RICE (1998, 2000)].

Given two tensors of the same rank, the inner product (symbolized by  $\langle \cdot, \cdot \rangle$ ) of them is just the sum of the products of all the corresponding elements, which is a number (the inner product of two tensors of different rank is not a number, but this is not an issue in this article). For example, the inner product of  $\mathbf{P}^3$  and  $\mathbf{D}^3$  is calculated as

$$\langle \mathbf{P}^3, \mathbf{D}^3 \rangle = \sum_i \sum_j \sum_k \mathbf{P}_{ijk}^3 \mathbf{D}_{ijk}^3. \quad (3)$$

The familiar example of the inner product of two vectors (tensors of rank one) is a special case of this. (What I

am calling the inner product is sometimes called the “ $n$ -fold inner product,” to signify that we are summing over all  $n$  indices.)

The outer product ( $\otimes$ ) of two tensors of rank  $r_1$  and  $r_2$  is another tensor of rank  $r_1 + r_2$ . The outer product is formed by multiplying each element of one by each element of the other. For example, the outer product of  $\mathbf{D}_1^1$  and  $\mathbf{D}_2^1$ , both of which have rank one, is a new tensor of rank two, which we can write as a matrix in which the  $ij$ th element is the product of the  $i$ th element of  $\mathbf{D}_1^1$  and the  $j$ th element of  $\mathbf{D}_2^1$ . If each trait is influenced by the same two underlying factors, then

$$\mathbf{D}_1^1 = \begin{bmatrix} \frac{\partial \phi_1}{\partial u_1} \\ \frac{\partial \phi_1}{\partial u_2} \end{bmatrix}, \mathbf{D}_2^1 = \begin{bmatrix} \frac{\partial \phi_2}{\partial u_1} \\ \frac{\partial \phi_2}{\partial u_2} \end{bmatrix} \Rightarrow \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 = \begin{bmatrix} \frac{\partial \phi_1}{\partial u_1} \frac{\partial \phi_2}{\partial u_1} & \frac{\partial \phi_1}{\partial u_1} \frac{\partial \phi_2}{\partial u_2} \\ \frac{\partial \phi_1}{\partial u_2} \frac{\partial \phi_2}{\partial u_1} & \frac{\partial \phi_1}{\partial u_2} \frac{\partial \phi_2}{\partial u_2} \end{bmatrix}. \quad (4)$$

The general rule for tensors of arbitrary rank is

$$\begin{aligned} (\mathbf{D}_1^n \otimes \mathbf{D}_2^m)_{x_1, \dots, x_n, y_1, \dots, y_m} &= (\mathbf{D}_1^n)_{x_1, \dots, x_n} (\mathbf{D}_2^m)_{y_1, \dots, y_m} \\ &= \frac{\partial^n \phi_1}{\partial u_{x_1} \dots \partial u_{x_n}} \frac{\partial^m \phi_2}{\partial u_{y_1} \dots \partial u_{y_m}}. \end{aligned} \quad (5)$$

The outer product is good for forming all combinations of elements, and the inner product is good for summing over multiples of these. Outer multiplication uses the same symbol ( $\otimes$ ) as the Kronecker product and in fact is identical to the Kronecker product when applied to vectors. The two operations are different, though, for higher-rank tensors. For example, the Kronecker product of two matrices is another matrix, while the outer product of two second-rank tensors (which can be written as matrices) is a fourth-rank tensor, which cannot be written as a matrix.

An underlying factor may be any measurable value, continuous or discrete, that influences the phenotypic traits of interest. The fact that the theory puts few restrictions on how these factors are defined and measured (see the discussion of coordinates below and in the APPENDIX) is one of the virtues of the phenotype landscape approach, since it allows us to simultaneously consider the expression of individual genes, the activities of complex enzymes, environmental factors, and any other quantity that is biologically relevant in the system of interest. Just what the underlying factors are is thus determined by the particular system under study and is not necessary for the derivation of general results. Nonetheless, it is often helpful to visualize something (gene sequence, enzyme, etc.) when working through the derivations below. Some areas of research that yield values that could be treated as underlying factors for characters of interest are models of metabolic processes based on enzymatic activity, quantitative trait locus analysis, and studies of gene expression rates.

The activities of different enzymes are often the main parameters in models of metabolic pathways. Such models, in which the trait being studied is the flux through a particular pathway, allow us to derive phenotype landscapes directly from the models (NIJHOUT 2002). In terms of the theory presented here, we know how to calculate the  $\mathbf{D}$  tensors for models involving enzymatic pathways. The  $\mathbf{D}$  tensors can also be estimated, in this case from empirical data, using quantitative trait locus (QTL) analysis. While quantitative trait loci are mechanistically ambiguous, being regions of a chromosome that somehow contribute to variation in a trait, they have the advantage that the degree to which each QTL contributes to a trait can be estimated, along with first- and second-order epistatic interactions.

Although metabolic models and QTL analyses allow us to identify the underlying factors for certain traits and to calculate the  $\mathbf{D}$  tensors, they yield no information about the distribution of underlying variation in a population, which defines the  $\mathbf{P}$  tensors. Gene expression rates, while less often incorporated into mathematical models of phenotype, have the advantage that we can estimate the actual population level distribution of expression rates using DNA microchip methods. The possibility of combining QTL analysis with DNA chip analysis (SCHADT *et al.* 2003) could potentially provide a way to fill in all the pieces.

#### COVARIANCE BETWEEN TRAITS

Two traits,  $\phi_1$  and  $\phi_2$ , correspond to two different phenotype landscapes over the set of all underlying factors that contribute to either trait (RICE 2000, 2002; WOLF *et al.* 2001). Each trait thus has its own set of  $\mathbf{D}$  tensors describing the local geometry of its landscape, with  $\mathbf{D}_i^j$  containing all of the  $i$ th derivatives of trait  $\phi_j$ . There is only one set of  $\mathbf{P}$  tensors, though, because the distribution of underlying factors is a property of the population as a whole. The APPENDIX shows that we can write the phenotypic covariance between the two traits as

$$\text{Cov}(\phi_1, \phi_2) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{i!j!} \langle \mathbf{P}^{i+j} - \mathbf{P}^i \otimes \mathbf{P}^j, \mathbf{D}_1^i \otimes \mathbf{D}_2^j \rangle. \quad (6)$$

The variance of either trait is found by calculating its covariance with itself; thus

$$\text{Var}(\phi_1) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{i!j!} \langle \mathbf{P}^{i+j} - \mathbf{P}^i \otimes \mathbf{P}^j, \mathbf{D}_1^i \otimes \mathbf{D}_1^j \rangle. \quad (7)$$

Equation 6 describes the phenotypic covariance between two traits as a function of the distribution of underlying factors and the manner in which those factors interact in development of the traits. If each of the underlying factors is a gene that is passed accurately to the next generation, then we are also calculating the genetic

covariance between the traits. Below, I consider the special case of Equation 6 appropriate to uncurved landscapes and then turn to the effects of curvature (*i.e.*, epistasis and dominance) on covariation between traits. [It is interesting to note that the form of the right-hand part of Equation 6—the inner product of a  $\mathbf{P}$  tensor with the outer product of a set of  $\mathbf{D}$  tensors—is similar to the form of the vectors defining the direction of evolution under selection (RICE 2002).]

**Uncurved landscapes:** If both phenotype landscapes are uncurved (completely additive), then  $\mathbf{D}^n$  is zero for all  $n > 1$ , so we need consider only the first term:

$$\text{Cov}(\phi_1, \phi_2) = \langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle. \quad (8)$$

This equation captures the contributions to phenotypic covariance of both covariance between the underlying factors (the off-diagonal elements of  $\mathbf{P}^2$ ) and pleiotropy. Figure 1 illustrates how a particular covariance between two traits may be due to either covariation of the underlying factors or pleiotropy. In the case where  $\phi_1$  is the value of the trait in the parents and  $\phi_2$  is the expected offspring value, then  $\langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle$  is the additive genetic variance for the trait (KOJIMA 1959).

Figure 1A shows the contour lines and gradient vectors ( $\mathbf{D}_a^1 = \nabla \phi_a$ ) for two traits defined by the functions

$$\begin{aligned} \phi_1 &= u_1 + \frac{1}{2}u_2 \\ \phi_2 &= \frac{1}{2}u_1 + u_2. \end{aligned} \quad (9)$$

Because the phenotype functions in Equation 9 are linear (second and higher derivatives are all zero) we need consider only  $\mathbf{P}^2$  (corresponding to  $\mathbf{D}^1 \otimes \mathbf{D}^1$ ), which is the same as the covariance matrix. If the two underlying factors are uncorrelated and have unit variances, then the relevant tensors are

$$\begin{aligned} \mathbf{D}_1^1 &= \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{D}_2^1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{P}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \mathbf{D}_1^1 \otimes \mathbf{D}_1^1 &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{D}_2^1 \otimes \mathbf{D}_2^1 = \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 2 & 1 \end{bmatrix}, \quad \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 2 \end{bmatrix} \end{aligned} \quad (10)$$

and, from Equations 6 and 7, we find

$$\begin{aligned} \text{Cov}(\phi_1, \phi_2) &= \langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle = \frac{1}{2} + \frac{1}{2} = 1 \\ \text{Var}(\phi_1) &= \langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_1^1 \rangle = 1 + \frac{1}{4} = 1.25 \\ \text{Var}(\phi_2) &= \langle \mathbf{P}^2, \mathbf{D}_2^1 \otimes \mathbf{D}_2^1 \rangle = \frac{1}{4} + 1 = 1.25. \end{aligned} \quad (11)$$

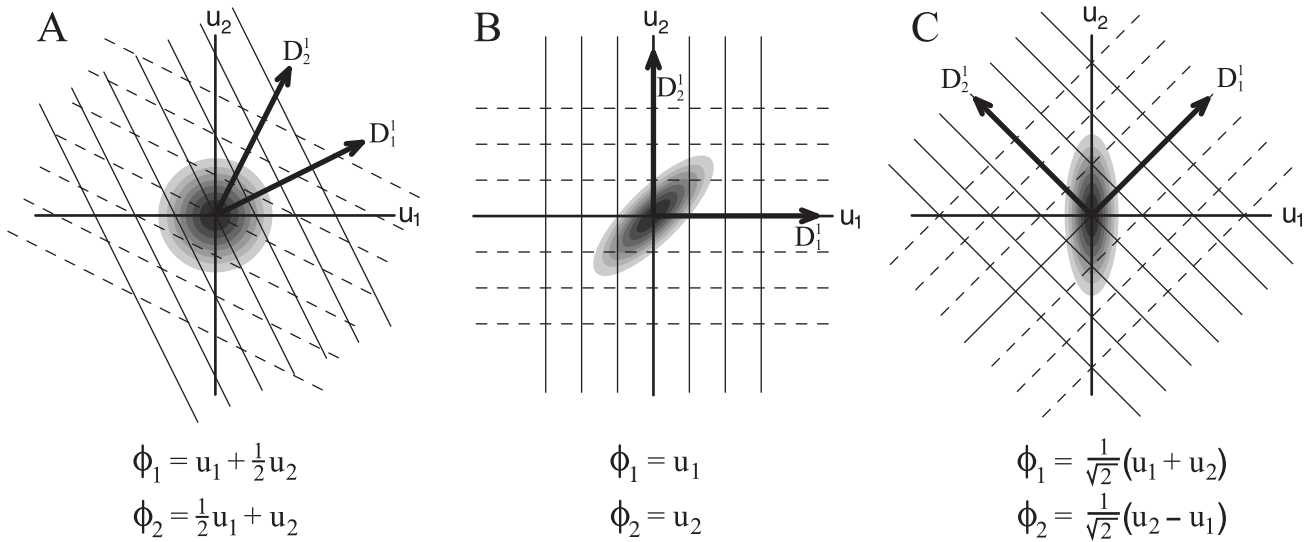


FIGURE 1.—Three situations that produce a correlation of 0.8 between traits  $\phi_1$  and  $\phi_2$ . The axes are the values of two underlying factors that contribute to the traits, the thin lines are contours of the phenotypic landscapes, the boldface arrows are the gradient vectors for the traits, and the shaded regions represent the distribution of underlying variation. (A) Landscapes corresponding to Equation 9. (B) Equation 12. (C) Equation 13.

So the correlation between traits  $\phi_1$  and  $\phi_2$  is 0.8 [since  $\text{Cor}(\phi_1, \phi_2) = \text{Cov}(\phi_1, \phi_2) \cdot (\text{Var}(\phi_1)\text{Var}(\phi_2))^{-1/2}$ ]. The case shown in Figure 1B yields the same correlation in a different way. Here, the covariance between the traits results from covariance between the underlying factors; the relevant tensors are

$$\mathbf{D}_1^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{D}_2^1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{P}^2 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

$$\mathbf{D}_1^1 \otimes \mathbf{D}_1^1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{D}_2^1 \otimes \mathbf{D}_2^1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (12)$$

Calculating the covariance and variances as above yields in this case  $\text{Cov}(\phi_1, \phi_2) = 0.8$  and  $\text{Var}(\phi_1) = \text{Var}(\phi_2) = 1$ , so the correlation is again 0.8.

In Figure 1A we would say that the correlation between the traits is due to pleiotropy, whereas in the case in Figure 1B we would say that the traits covary because of nonindependence of the underlying factors. It is important to note, though, that terms like “pleiotropy” are not necessarily inherent properties of the system, but rather are functions of how we choose to draw our coordinates. Figure 1C shows a system with exactly the same geometry as that in Figure 1B, the only difference being that everything is rotated by  $45^\circ$ . In this case, both underlying factors influence both traits, so we would say that there is pleiotropy, even though the geometry is exactly the same as in Figure 1B.

The relevant tensors in the case shown in Figure 1C are

$$\mathbf{D}_1^1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{D}_2^1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathbf{P}^2 = \begin{bmatrix} 0.2 & 0 \\ 0 & 1.8 \end{bmatrix}$$

$$\mathbf{D}_1^1 \otimes \mathbf{D}_1^1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{D}_2^1 \otimes \mathbf{D}_2^1 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}. \quad (13)$$

Thus, although rotating the coordinate system causes the elements of each tensor to change, the relations between them, including the value of  $(\mathbf{P}, \mathbf{D} \otimes \mathbf{D})$ , do not change. This invariance under rotation is one of the defining characteristics of tensors. For our purposes, it means that we have some leeway in how we choose to define the underlying factors; so long as we choose factors that span the space of underlying variation, we can choose a coordinate system that captures what we know about the biology involved. Alternately, it is often useful to rotate the coordinates so as to simplify the calculations by making as many terms equal to zero as possible (compare the analyses for Figure 1, B and C, above). The fact that the  $\mathbf{P}$  and  $\mathbf{D}$  tensors exist independently of any particular coordinate system is also important when we recognize that our identification of underlying factors, both genetic and otherwise, is somewhat arbitrary.

Changes in the coordinate system that are more complex than simple rotation are discussed in the APPENDIX. Although in this article I discuss tensors in terms of their elements in some given coordinate system, we can think of them as being independent of the way in which we choose to measure things. For example,  $\mathbf{D}^2$  captures the phenotypic consequences of second-order interactions

among the underlying factors; if we change the way in which we measure those factors, the elements of  $\mathbf{D}^2$  change, but its relation to the other tensors does not.

We can further simplify Equation 6 in the special case in which the underlying factors are uncorrelated with one another and are all measured in units of standard deviations (*i.e.*, scaled so that the variances are all equal to 1); then  $\mathbf{P}^2$  is just the identity matrix and  $\langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle$  is just the sum of the diagonal elements of  $\mathbf{D}_1^1 \otimes \mathbf{D}_2^1$ , which is  $\sum \partial \phi_1 / \partial u_i \cdot \partial \phi_2 / \partial u_i$ . This is the same as the inner product of the two gradient vectors, so for this special case (zero covariances and variances equal to one) we have

$$\text{Cov}(\phi_1, \phi_2) = \langle \nabla \phi_1, \nabla \phi_2 \rangle. \quad (14)$$

A standard result from the definition of the inner product of two vectors is

$$\langle \nabla \phi_1, \nabla \phi_2 \rangle = \|\nabla \phi_1\| \cdot \|\nabla \phi_2\| \cdot \text{Cos}(\theta), \quad (15)$$

where  $\|\nabla \phi_i\|$  is the length of the vector  $\nabla \phi_i$ , and  $\theta$  is the angle between the two gradient vectors. Under the assumption of unit variances and no covariances between the underlying factors, the variances of the traits are given by WOLF *et al.* (2001):

$$\text{Var}(\phi_i) = \|\nabla \phi_i\|^2. \quad (16)$$

Combining Equations 14, 15, and 16, we can calculate the correlation coefficient for the two traits as

$$\text{Cor}(\phi_1, \phi_2) = \text{Cos}(\theta). \quad (17)$$

Thus, in an additive system with unit variances and no covariances between the underlying factors, the correlation between the two traits is exactly the cosine of the angle between their gradient vectors.

Interestingly, there are cases in which the covariance between the underlying factors has no effect on the covariance between the traits. If one of the gradient vectors is a reflection of the other in one of the axes, then the covariance between the two traits is determined by the angle between their gradient vectors and by the variances (but not covariances) of the underlying factors. This is illustrated in Figure 2. This independence of phenotypic covariance and underlying covariance results because  $\mathbf{P}^2$  is symmetrical ( $\mathbf{P}_{ij}^2 = \mathbf{P}_{ji}^2$ ) but  $(\mathbf{D}_1^1 \otimes \mathbf{D}_2^1)$  is not. Whenever  $(\mathbf{D}_1^1 \otimes \mathbf{D}_2^1)_{ij} = -(\mathbf{D}_1^1 \otimes \mathbf{D}_2^1)_{ji}$ , then the two terms containing  $\mathbf{P}_{ij}^2$  cancel one another out, so  $\mathbf{P}_{ij}^2$  contributes nothing to  $\text{Cov}(\phi_1, \phi_2)$ . In such a case, selection for integration of the two phenotypic traits will not lead to covariance of the underlying factors (such as through gametic phase disequilibrium), although it may lead to change in their variances.

**Curved landscapes:** Consider first two quadratic landscapes. Using the fact that  $\mathbf{P}^1 = 0$  (since we are using central moments), Equation 6 yields

$$\text{Cov}(\phi_1, \phi_2) = \langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle + \frac{1}{2} \langle \mathbf{P}^3, \mathbf{D}_1^1 \otimes \mathbf{D}_2^2 \rangle + \frac{1}{2} \langle \mathbf{P}^3, \mathbf{D}_1^2 \otimes \mathbf{D}_2^1 \rangle$$

$$+ \frac{1}{4} \langle \mathbf{P}^4 - \mathbf{P}^2 \otimes \mathbf{P}^2, \mathbf{D}_1^2 \otimes \mathbf{D}_2^2 \rangle. \quad (18)$$

Equation 18 shows that curvature of either or both of the phenotype landscapes (captured by the  $\mathbf{D}^2$  terms) can contribute to the phenotypic covariance. To see how this occurs, consider the case of quadratic phenotype landscapes and a multivariate normal distribution of underlying variation with no covariances between the underlying factors and equal variances designated by  $\sigma^2$ . Under these circumstances  $\mathbf{P}^3 = 0$ ,  $\mathbf{P}_{1111}^4 = \mathbf{P}_{2222}^4 = 3\sigma^4$ ,  $\mathbf{P}_{ijkm}^4 = \sigma^4$  when the sequence *ijkm* contains exactly two 1's and two 2's, and  $\mathbf{P}_{\text{else}}^4 = 0$ . Focusing on the fourth term on the right-hand side of Equation 18, we find

$$\begin{aligned} (\mathbf{P}^4 - \mathbf{P}^2 \otimes \mathbf{P}^2)_{1111 \& 2222} &= 2\sigma^4 \\ (\mathbf{P}^4 - \mathbf{P}^2 \otimes \mathbf{P}^2)_{1221 \& 2112 \& 1212 \& 2121} &= \sigma^4 \\ (\mathbf{P}^4 - \mathbf{P}^2 \otimes \mathbf{P}^2)_{\text{else}} &= 0. \end{aligned} \quad (19)$$

The condition for epistasis to contribute to phenotypic covariance given these assumptions is thus

$$\frac{\partial^2 \phi_1}{\partial u_1^2} \frac{\partial^2 \phi_2}{\partial u_1^2} + \frac{\partial^2 \phi_1}{\partial u_2^2} \frac{\partial^2 \phi_2}{\partial u_2^2} + 2 \frac{\partial^2 \phi_1}{\partial u_1 \partial u_2} \frac{\partial^2 \phi_2}{\partial u_1 \partial u_2} \neq 0. \quad (20)$$

Figure 3 shows a case that meets this condition. Even when the gradient vectors are at right angles and there is no covariance between underlying factors, epistatic effects produce a covariance between the phenotypic traits.

Phenotypic or genetic covariance induced by epistatic effects differs in two important ways from covariance resulting from additive pleiotropic effects or covariance between underlying factors. First, in cases such as the one in Figure 3, the correlation between the traits is a function of the total amount of underlying variation. Uniformly reducing the variances in both underlying factors in Figure 3 would reduce the correlation between  $\phi_1$  and  $\phi_2$ , because the landscapes would look more nearly flat over the range of variation present in the population. This is not the case in the examples shown in Figure 1, where increasing or decreasing the amount of underlying variation, without changing the shape of the distribution, has no effect on the correlation between the phenotypic traits.

Second, the association between  $\phi_1$  and  $\phi_2$  in Figure 3 is asymmetrical. Curvature of the landscapes makes it easier to simultaneously increase both  $\phi_1$  and  $\phi_2$  than it would if there were no epistasis, but it is actually harder to simultaneously decrease both traits, since the proportion of the distribution of variation that lies in the region in which both traits are smaller is less than it would be with the same gradient vectors on uncurved landscapes.

A number of studies, including BELL and BURRIS (1973) with *Tribolium*, RUTLEDGE *et al.* (1973) with mice, and NORDSKOG (1977) with chickens, have observed an asymmetry in selection experiments on correlated characters [*e.g.*, body weight and tail length in mice (RUTLEDGE *et*

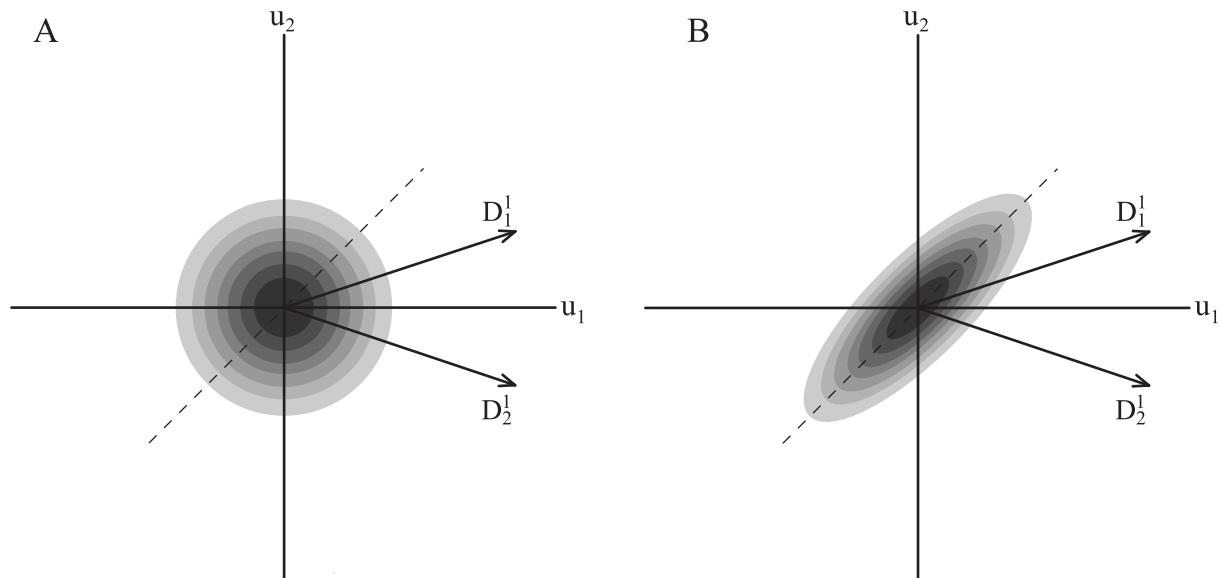


FIGURE 2.—A case in which covariation between the underlying factors does not influence covariation between the traits. In each case the correlation between  $\phi_1$  and  $\phi_2$  is 0.8. The shaded regions represent the distribution of underlying variation in a population. In A there is no covariance between the underlying factors. In B, the underlying factors covary.

*al.* 1973)] in which the response to selection over many generations is significantly greater for one kind of antagonistic selection (*e.g.*, small body and long tail) than for the other (*e.g.*, large body and short tail). This sort of pattern is not predicted by models based on only additive gene effects, but follows from a number of different scenarios involving epistasis.

Classical quantitative genetics assumes locally uncurved landscapes since this means that the local geometry is well approximated by an additive model. A common way to test this assumption is to partition genetic

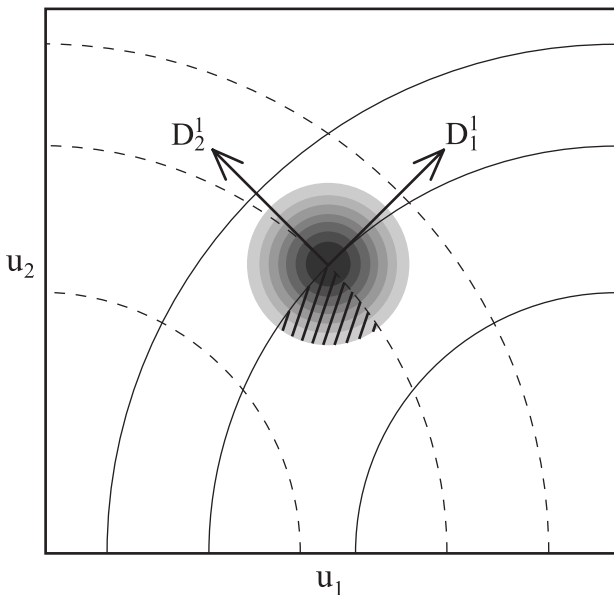


FIGURE 3.—A case in which covariation between two traits is a result of epistasis only. The gradient vectors are at right angles. The shaded circles represent the distribution of underlying variation. The hatched region is the subset of that distribution corresponding to both traits being smaller than the population mean.

variance into additive and nonadditive components and then check the magnitude of the nonadditive components (LYNCH and WALSH 1998). A number of authors have pointed out that this approach can easily underestimate the significance of epistatic effects (KEIGHTLEY 1989; LYNCH and WALSH 1998). Selection experiments in which the results are compared quantitatively with the predictions of an additive model provide a more sensitive test of whether epistasis is strong enough to influence evolution, and the observation of a number of cases in which the response to antagonistic selection is asymmetrical (NORDSKOG 1977) suggests that many traits do not have locally uncurved landscapes.

Figure 4 shows a hypothetical case in which two underlying factors influence two traits. Factor  $u_2$  has a linear positive effect on both traits, while factor  $u_1$  has a positive effect on one trait and a negative effect on the other. The key is that the contribution of  $u_1$  is highly nonlinear, so that it has little effect on either trait when it is small and then rapidly increases in importance as its value increases. This creates a situation in which a relatively small part of the distribution of underlying factors lies in the region that produces an increased value of  $\phi_1$  and a decreased value of  $\phi_2$ . Furthermore, moving in the direction necessary to increase  $\phi_1$  while decreasing  $\phi_2$  moves the population into a region in which the two traits are even more highly correlated. Figure 4 shows a two-dimensional case purely for the sake of drawability; the same kind of result can occur with any number of underlying factors.

This sort of effect is apparent in the “resource allocation” models studied by HOULE (1991; see also DE JONG and VAN NOORDWIJK 1992). Here, the development of two phenotypic traits involves the same resources, with

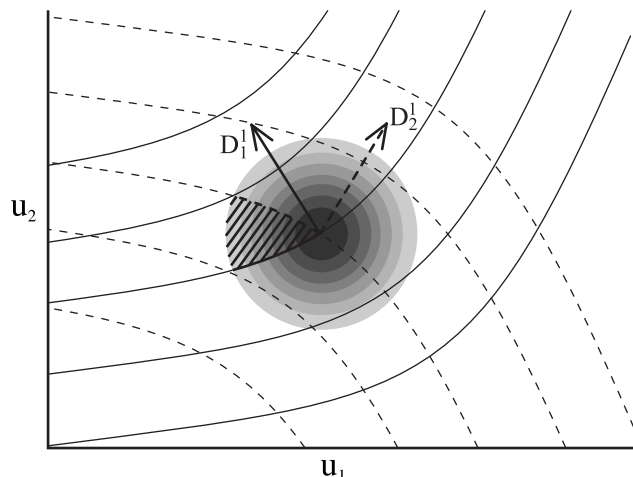


FIGURE 4.—A case in which the covariation induced by epistasis is asymmetrical, such that the response to selection for increasing  $\phi_1$  while decreasing  $\phi_2$  (selecting values in the hatched region) should be greater than that to selection for decreasing  $\phi_2$  while increasing  $\phi_1$ .

the values of the traits being determined by the rate at which resources are acquired and how they are distributed. Letting  $u_1$  be the rate at which the resource is acquired and  $u_2$  be the proportion of the resource allocated to trait  $\phi_1$ , the two traits are described by

$$\begin{aligned}\phi_1 &= u_1 u_2 \\ \phi_2 &= u_1(1 - u_2).\end{aligned}\quad (21)$$

Here,  $u_1$  and  $u_2$  are themselves quantitative traits that may be influenced by many genes. For simplicity, these are assumed to have heritability 1 and be uncorrelated. The two phenotype landscapes for these traits are shown in Figure 5.

The relevant  $\mathbf{D}$  tensors for the traits defined in Equations 21 are

$$\mathbf{D}_1^1 = \begin{bmatrix} u_2 \\ u_1 \end{bmatrix}, \quad \mathbf{D}_2^1 = \begin{bmatrix} 1 - u_2 \\ -u_1 \end{bmatrix}, \quad \mathbf{D}_1^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{D}_2^2 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.\quad (22)$$

From these we find

$$\begin{aligned}(\mathbf{D}_1^2 \otimes \mathbf{D}_2^2)_{1212&1221&2112&2121} &= -1 \\ (\mathbf{D}_1^2 \otimes \mathbf{D}_2^2)_{\text{else}} &= 0.\end{aligned}$$

Given the assumption that  $\mathbf{u}_1$  and  $\mathbf{u}_2$  have zero covariance, we know that  $\mathbf{P}_{ij}^2 = 0$  for  $i \neq j$ , so we have

$$(\mathbf{P}^2 \otimes \mathbf{P}^2)_{1212&1221&2112&2121} = 0.$$

Thus,  $\langle \mathbf{P}^2 \otimes \mathbf{P}^2, \mathbf{D}_1^2 \otimes \mathbf{D}_2^2 \rangle = 0$  and, given the assumption that  $\mathbf{P}^3 = 0$  (symmetrical distribution), we can write

$$\text{Cov}(\phi_1, \phi_2) = \langle \mathbf{P}^2, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle + \frac{1}{4} \langle \mathbf{P}^4, \mathbf{D}_1^2 \otimes \mathbf{D}_2^2 \rangle.\quad (23)$$

Using the fact that the only nonzero elements of  $\mathbf{P}^4$  are

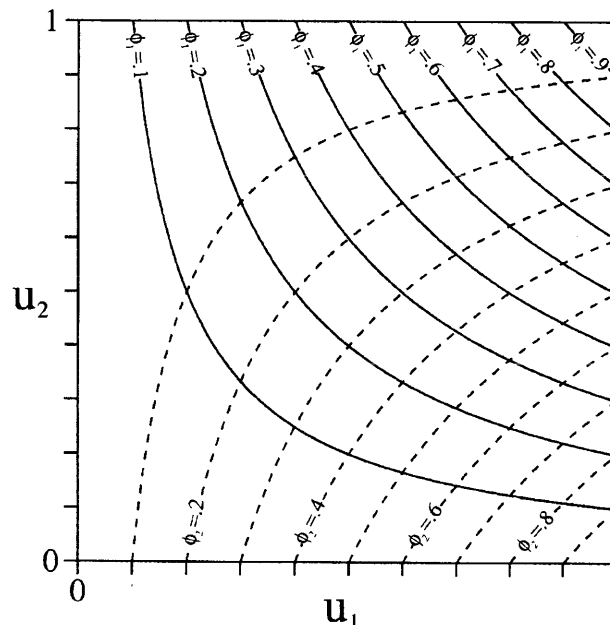


FIGURE 5.—Phenotype landscapes corresponding to Equation 21. The solid lines are contours of the  $\phi_1$  landscape and the dashed lines are contours of the  $\phi_2$  landscape.

$\mathbf{P}_{1212&1221&2112&2121}^4 = \sigma_1^2 \sigma_2^2$  and the values of the tensors in Equation 22 we find

$$\text{Cov}(\phi_1, \phi_2) = \sigma_1^2 u_2(1 - u_2) - \sigma_2^2 u_1^2 - \sigma_1^2 \sigma_2^2.\quad (24)$$

The sum of the first two terms on the right side of Equation 24 [ $\sigma_1^2 u_2(1 - u_2) - \sigma_2^2 u_1^2$ ] is the covariance between the additive approximations to Equations 21 and is thus the “additive genetic covariance” between the traits (DE JONG and VAN NOORDWIJK 1992; ROFF 1997). The third term on the right side of Equation 24,  $\sigma_1^2 \sigma_2^2$ , is the contribution of nonadditive interactions to covariance between the traits. In traditional nomenclature this is the “additive-by-additive” epistatic variance (KOJIMA 1959; DE JONG and VAN NOORDWIJK 1992).

Equation 24 has been derived elsewhere and follows from the properties of the multivariate normal distribution, so using the method presented here may seem like overkill. In fact, quite a bit more information can be gleaned from the  $\mathbf{D}$  tensors in Equations 22. For example, relaxing the assumption of multivariate normality requires that we consider the third moments (measuring asymmetry) of the distribution of underlying factors. Simply by examining the  $\mathbf{D}$  tensors in Equation 22, we can conclude that introducing skewness in the distribution of either  $u_1$  or  $u_2$  will have no effect on the resulting covariance between  $\phi_1$  and  $\phi_2$ , but changing the mixed third moments (e.g.,  $E[x_1 x_2^2]$ ) will change the value of  $\text{Cov}(\phi_1, \phi_2)$ .

Recall that the third moments influence covariance between the traits through the terms  $\langle \mathbf{P}^3, \mathbf{D}_1^1 \otimes \mathbf{D}_2^1 \rangle$  (see Equation 18). Since  $(\mathbf{D}_1^1 \otimes \mathbf{D}_2^1)_{111} = (\mathbf{D}_1^1)_1 \cdot (\mathbf{D}_2^1)_{11}$ , we im-

mediately see that because the diagonal elements of both  $\mathbf{D}_1^2$  and  $\mathbf{D}_2^2$  are all zero,  $(\mathbf{D}_1^1 \otimes \mathbf{D}_2^2)_{111}$  and  $(\mathbf{D}_1^1 \otimes \mathbf{D}_2^2)_{222}$  must both be zero, as must the corresponding elements of  $\mathbf{D}_2^1 \otimes \mathbf{D}_1^2$ . We thus know that whatever the values of  $\mathbf{P}_{111}$  and  $\mathbf{P}_{222}$ , which measure the skewness in the  $u_1$  and  $u_2$  directions, respectively, they have no influence on the covariance between  $\phi_1$  and  $\phi_2$ , since they are both multiplied by terms that equal zero. This is relevant to interpreting the generality of the model; because  $u_1$  has a lower bound at zero and no upper bound, it is reasonable to expect that the distribution of values of  $u_1$  would be positively skewed. An analogous argument shows that the mixed third moments, such as  $\mathbf{P}_{112}$  and  $\mathbf{P}_{122}$ , do contribute to the covariance between the traits. These moments measure, roughly, the degree to which the conditional variance of one underlying factor is a function of the value of the other factor (there is an example in the final section below).

PARENT-OFFSPRING COVARIANCE AND HERITABILITY

One interesting application of this theory is to the calculation of parent-offspring covariance. Here, instead of considering two different traits in the same generation, we calculate the covariance between a trait in the parents,  $\phi_p$ , and the same trait among their offspring,  $\phi_o$ . Dividing this covariance by the variance in parental phenotype yields the regression of offspring on parents,  $\beta_{o,p}$ , which I treat as identical to heritability. In sexually reproducing diploid organisms, we can use the midparent phenotypic value as  $\phi_p$  without changing our calculations. [A number of different definitions of heritability exist in the literature (JACQUARD 1983); I focus on the regression of offspring on parents and the corresponding offspring-parent covariance, since this is most directly related to the response to selection.]

Consider first the simplest case, in which phenotype is an additive function of a heritable underlying factor,  $u_1$ , and an environmental factor. We assume that the expected value of  $u_1$  is the same in offspring as in their parents, but that the environment experienced by the offspring is (potentially) independent of that of their parents. In such a case, we really have two different environmental factors: the environment experienced by the parents,  $u_2$ , and the environment experienced by their offspring,  $u_3$ . We can now write the phenotype of the parents,  $\phi_p$ , and that of their offspring,  $\phi_o$ , as

$$\begin{aligned} \phi_p &= \alpha u_1 + \epsilon u_2 \\ \phi_o &= \alpha u_1 + \epsilon u_3. \end{aligned} \tag{25}$$

To find the parent-offspring covariance we note that the relevant  $\mathbf{D}$  tensors are

$$\mathbf{D}_p^1 = \begin{bmatrix} \alpha \\ \epsilon \\ 0 \end{bmatrix}, \mathbf{D}_o^1 = \begin{bmatrix} \alpha \\ 0 \\ \epsilon \end{bmatrix}, \mathbf{D}_p^1 \otimes \mathbf{D}_o^1 = \begin{bmatrix} \alpha^2 & 0 & \alpha\epsilon \\ \alpha\epsilon & 0 & \epsilon^2 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{D}_p^1 \otimes \mathbf{D}_p^1 = \begin{bmatrix} \alpha^2 & \alpha\epsilon & 0 \\ \alpha\epsilon & \epsilon^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{26}$$

from which, using Equations 6 and 7, we find

$$\begin{aligned} \text{Cov}(\phi_o, \phi_p) &= \alpha^2 P_{11} + \alpha\epsilon P_{13} + \alpha\epsilon P_{21} + \epsilon^2 P_{23} \\ \text{Var}(\phi_p) &= \alpha^2 P_{11} + \alpha\epsilon P_{12} + \alpha\epsilon P_{21} + \epsilon^2 P_{22}. \end{aligned} \tag{27}$$

The regression of offspring on parents,  $\beta_{o,p}$ , is then

$$\beta_{o,p} = \frac{\text{Cov}(\phi_o, \phi_p)}{\text{Var}(\phi_p)} = \frac{\alpha^2 P_{11} + \alpha\epsilon P_{13} + \alpha\epsilon P_{21} + \epsilon^2 P_{23}}{\alpha^2 P_{11} + \alpha\epsilon P_{12} + \alpha\epsilon P_{21} + \epsilon^2 P_{22}}. \tag{28}$$

The terms  $P_{12}$ ,  $P_{13}$ , and  $P_{23}$  are, respectively, the covariance between the parent's genotype and the parent's environment ( $P_{12} = P_{21}$ ), the covariance between the parent's genotype and the offspring's environment ( $P_{13}$ ), and the covariance between the parent's environment and the offspring's environment ( $P_{23}$ ). If all of these covariances are zero, then we have

$$\beta_{o,p} = \frac{\alpha^2 P_{11}}{\alpha^2 P_{11} + \epsilon^2 P_{22}}, \tag{29}$$

which corresponds to the standard quantitative genetics notion of heritability, with the denominator being the total phenotypic variance among the parents and the numerator being the component of that variance that is due to heritable factors (the "additive genetic variance"). Note, though, that when we use the full Equation 28 the numerator contains terms that do not appear in the denominator, so we cannot represent the regression of offspring on parents in terms of the ratio of additive genetic variance to phenotypic variance, since this would require that the additive genetic variance be either negative or greater than the phenotypic variance.

In the linear case of Equation 25, the regression of offspring phenotype on parent phenotype is not a function of the values of the underlying factors. This is not the case, though, when there are nonlinear effects, such as epistasis or genotype-by-environment interactions. As an example, consider a trait in the parental generation,  $\phi_p$ , influenced by two genetic factors,  $u_1$  and  $u_2$ , and the environment. For simplicity, we assume that  $u_1$  and  $u_2$  are transmitted without modification to the offspring (*i.e.*, assume asexual reproduction). As before, we have two environmental factors, the value of the environment experienced by the parents,  $u_3$ , and the environmental value experienced by their offspring,  $u_4$ :

$$\begin{aligned} \phi_p &= u_1^2 + u_2 u_3 \\ \phi_o &= u_1^2 + u_2 u_4. \end{aligned} \tag{30}$$

Note that the two equations are the same except that  $\phi_p$  contains  $u_3$ , the parent's environment, where  $\phi_o$  contains  $u_4$ , the offspring's environment. The first rank  $\mathbf{D}$  tensors are



$$\mathbf{D}_p^1 = \begin{bmatrix} 2u_1 \\ u_3 \\ u_2 \\ 0 \end{bmatrix}, \quad \mathbf{D}_o^1 = \begin{bmatrix} 2u_1 \\ u_4 \\ 0 \\ u_2 \end{bmatrix}. \quad (31)$$

Combining these according to Equation 4 yields

$$\mathbf{D}_p^1 \otimes \mathbf{D}_o^1 = \begin{bmatrix} 4u_1^2 & 2u_1u_4 & 0 & 2u_1u_2 \\ 2u_1u_3 & u_3u_4 & 0 & u_2u_3 \\ 2u_1u_2 & u_2u_4 & 0 & u_2^2 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{D}_p^1 \otimes \mathbf{D}_p^1 = \begin{bmatrix} 4u_1^2 & 2u_1u_3 & 2u_1u_2 & 0 \\ 2u_1u_3 & u_3^2 & u_2u_3 & 0 \\ 2u_1u_2 & u_2u_3 & u_2^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (32)$$

The second rank  $\mathbf{D}$  tensors are

$$\mathbf{D}_p^2 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{D}_o^2 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (33)$$

Although the outer product of these is a four-dimensional array, most of its elements are zero, and we can quickly find the nonzero elements as

$$\begin{aligned} (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{1111} &= 4 \\ (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{1124} &= (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{1142} = (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{2311} = (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{3211} = 2 \\ (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{2324} &= (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{2342} = (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{3224} = (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{3242} = 1 \\ (\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{\text{else}} &= 0. \end{aligned} \quad (34)$$

Each of these will be multiplied by the corresponding element of  $\mathbf{P}^4$ . We thus know that we can ignore  $\mathbf{P}_{1122}$ , for example, because  $(\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{1122} = 0$ .

Note that, except for  $(\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)_{1111}$ , all nonzero elements of  $(\mathbf{D}_p^2 \otimes \mathbf{D}_o^2)$  have subscripts with two numbers the same and one each of two other numbers (*e.g.*, 3242, with two 2's and a 3 and a 4). These will be multiplied by elements of the  $\mathbf{P}^4$  tensor that are moments of the form  $E[x_i^2 x_j x_k]$  for  $i \neq j \neq k$ . Fourth-order moments of this form are zero if there are no covariances between the underlying factors. Thus, for a normal distribution with all covariances equal to zero, we need consider only the term that multiplies  $\mathbf{P}_{1111}$ , which is 4. Applying these results to Equations 6 and 7 yields

$$\beta_{o,p} = \frac{\text{Cov}(\phi_o, \phi_p)}{\text{Var}(\phi_p)} = \frac{4u_1^2 \mathbf{P}_{11} + u_3 u_4 \mathbf{P}_{22} + \mathbf{P}_{1111}}{4u_1^2 \mathbf{P}_{11} + u_3^2 \mathbf{P}_{22} + u_2^2 \mathbf{P}_{33} + \mathbf{P}_{1111}}. \quad (35)$$

Figure 6 shows a slice of the phenotype landscape for the trait defined by Equation 30; the full landscape is four-dimensional; this slice is produced by fixing the values of the environmental factors, here  $u_3 = u_4 = 1$ . In Figure 6,  $\mathbf{P}_{11} = \mathbf{P}_{22} = \mathbf{P}_{33} = 1$  and  $\mathbf{P}_{1111} = 3$  (which is the fourth moment for a normal distribution with unit variance). Figure 6 also shows the heritability (the

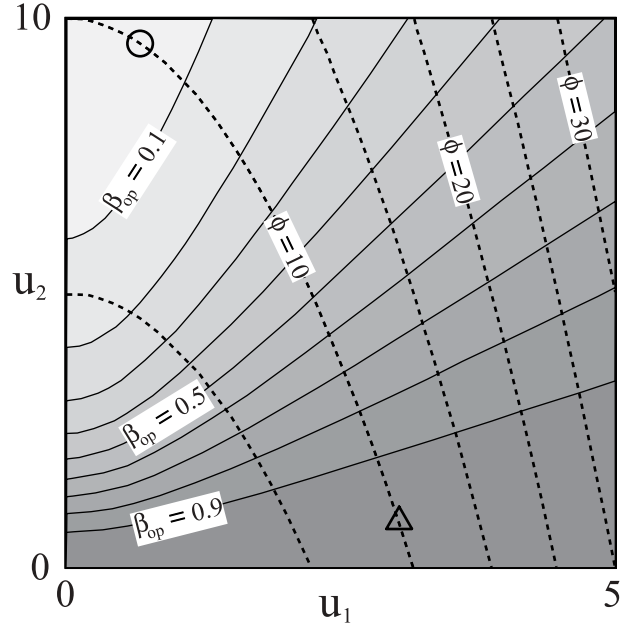


FIGURE 6.—Phenotype landscape (dashed contour lines) for the parental trait defined in Equation 30 superimposed on the offspring-parent regression (heritability) for the trait, given by Equation 35 (shading). Darker shading indicates higher heritability for the trait.

regression of offspring phenotype on parent phenotype) for the trait  $\phi$ , indicated by shading. Moving along the  $\phi = 10$  curve from the circle to the triangle produces no change in the mean phenotype, but causes the heritability to increase from nearly zero to nearly one.

Thus, when the underlying factors, both genetic and environmental, contribute nonadditively to a trait, then change in the values of those underlying factors tends to lead to change in the trait's heritability. (If all underlying factors contribute additively to the trait, then heritability changes only as a result of change in the elements of the  $\mathbf{P}$  tensors.)

This provides one way to visualize genetic assimilation. Moving along the dashed contour from the circle to the triangle in Figure 6 moves a population from a point where variation in the trait is largely determined by environmental variation to a point where most variation is heritable. This phenomenon may also play a role in the evolution of novel phenotypic traits. Think of a trait as a function of some set of underlying factors. Even if selection acts on the function (*i.e.*, fitness is determined by the value of the function), there will be no subsequent evolution if it has very low heritability (*e.g.*, the circle in Figure 6). If the distribution of underlying factors changes (for whatever reason) in such a way that the population comes to inhabit a region of the space of underlying factors in which the function in question is heritable (*e.g.*, the triangle in Figure 6), then selection can now act to modify the value of that

function, which now behaves in an evolutionary sense like a distinct trait.

### SELECTION DIFFERENTIALS

This article is concerned primarily with the statics, rather than the dynamics, of phenotypes, meaning that we are concerned with concepts such as genetic and phenotypic covariance, heritability, and the like, which are relevant to evolution but can be defined independently of any particular evolutionary process. I have thus made no formal reference to fitness. There is a link to selection theory, though, since the selection differential for a trait (the change due only to selection, prior to any changes introduced by recombination, etc.) is equal to the covariance between fitness and the trait value, divided by mean population fitness (ROBERTSON 1966; PRICE 1970). We can thus calculate the selection differential for the trait  $\phi$ , denoted  $S_\phi$ , by substituting fitness for one of the phenotypic traits in Equation 6:

$$S_\phi = \frac{1}{\bar{w}} \text{Cov}(w, \phi) = \frac{1}{\bar{w}} \sum_i \sum_j \frac{1}{i!j!} \langle \mathbf{P}^{i+j}, \mathbf{D}_w^i \otimes \mathbf{D}_\phi^j \rangle. \quad (36)$$

Equation 36 gives the effects of selection on a trait in terms of a fitness landscape and a phenotype landscape. Note, though, that the fitness landscape is over the space of underlying factors, rather than over the space of phenotypic traits, as in most quantitative genetic models (e.g., LANDE 1979). Note that Equation 36 gives the selection differential for the trait,  $\phi$ . It is thus different from the selection differential for the set of underlying factors derived elsewhere (RICE 2002).

### OTHER DEVELOPMENTAL ASSOCIATIONS BETWEEN TRAITS

Covariance between traits is of interest primarily because it influences the joint evolution of the traits, identifying cases in which directional selection on one trait is expected to produce a change in the mean value of both traits (LANDE 1979). In a completely additive model, in which all phenotype landscapes are uncurved, this is the only kind of developmental association that could influence joint evolution of the traits. When non-additive interactions are allowed between underlying factors, though, other kinds of relationships become possible, because moving in the space of underlying factors can change other moments of the phenotype distribution, not just the mean.

We can think of correlated evolution as being a case in which changing the first moment (the mean) of one trait leads to a change in the first moment of the other trait. With epistasis, it is possible to have a situation in which changing the  $n$ th moment of one trait leads to a change in the  $m$ th moment of another trait or to a change in the joint moments of a set of other traits.

I shall say that two phenotypic traits are *developmentally*

*entangled* if, because of developmental associations, changing some moment of the distribution of one trait leads to a change in some moment of the distribution of the other trait. Genetic covariance resulting from pleiotropy is one sort of developmental entanglement; below I give a couple of examples of other cases.

Consider two traits with phenotype landscapes defined by

$$\begin{aligned} \phi_1 &= \frac{1}{2}(u_1^2 - u_2^2) \\ \phi_2 &= u_1 + u_2. \end{aligned} \quad (37)$$

The contour maps of these are shown in Figure 7. The relevant  $\mathbf{D}$  tensors are

$$\mathbf{D}_1^1 = \nabla \phi_1 = \begin{bmatrix} u_1 \\ -u_2 \end{bmatrix}, \quad \mathbf{D}_1^2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{D}_2^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{D}_2^2 = [0]. \quad (38)$$

If the joint distribution of  $u_1$  and  $u_2$  is uncorrelated with equal variances, then the traits  $\phi_1$  and  $\phi_2$  are also uncorrelated (from Equation 6, using the fact that  $u_1 = u_2$  at this point). Thus, at least locally around this point, the two traits should evolve independently under directional selection. The two traits are not, however, independent under other kinds of selection.

As shown in RICE (1998), stabilizing selection on trait  $\phi_1$  should, all else held equal, move the population along the optimal contour toward a region of lower slope (as indicated by greater spacing of the contour lines). The direction of maximal decrease in slope for  $\phi_1$  is given by the vector  $-\mathbf{D}_1^2 \nabla \phi_1$  (RICE 1998). At the point  $u_1 = u_2 = 2$ , this vector points in the exact opposite direction of  $\nabla \phi_2$ . Thus, stabilizing selection on trait  $\phi_1$  leads to directional change of trait  $\phi_2$ , although the two traits are uncorrelated with one another.

We could not identify such a system by looking at just the covariance between the traits. This does not mean that there is no telltale signature, only that other moments of the joint distribution of the traits need to be considered. In the example above, a symmetrical distribution of underlying factors would produce an asymmetrical distribution of the traits, corresponding to a nonzero value of the mixed third moment  $E[x_1 x_2^2]$ , as shown in Figure 8.

The resource allocation model shown in Figure 5 also exhibits entanglement of the traits other than mere covariance. In Figure 5, it is not possible to change the degree of canalization of  $\phi_1$  without also changing the mean of  $\phi_2$ .

Any number of traits can potentially be entangled with one another. For example, consider the following system:

$$\begin{aligned} \phi_1 &= u_1 u_2 \\ \phi_2 &= u_1 u_3 \\ \phi_3 &= u_2 u_3. \end{aligned} \quad (39)$$

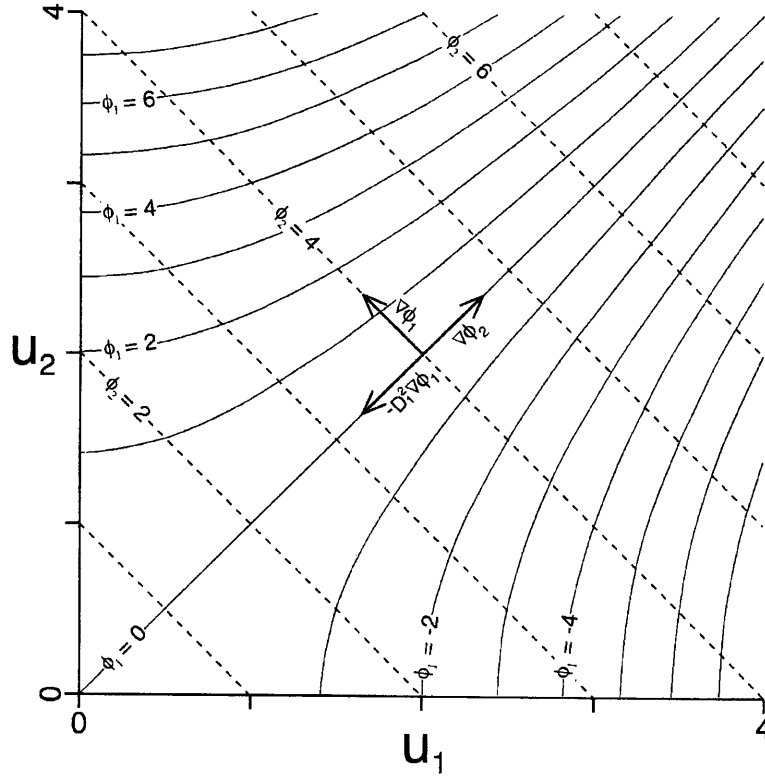


FIGURE 7.—Case in which stabilizing selection on one trait should lead to change in the mean of another trait. The solid contours correspond to the phenotype landscape of trait  $\phi_1$ . The dashed contours are for trait  $\phi_2$ . Stabilizing (canalizing) selection on trait  $\phi_1$  will move the population in the direction of  $-\mathbf{D}_1^i \nabla \phi_1$ .

In this system, there are three phenotype landscapes in a space of three underlying factors, which means that we would need four dimensions to draw a picture of it. While this hinders illustration, it does not get in the way of analysis.

The gradient vectors for these three traits are

$$\mathbf{D}_1^i = \nabla \phi_1 = \begin{bmatrix} u_2 \\ u_1 \\ 0 \end{bmatrix}, \quad \mathbf{D}_2^i = \nabla \phi_2 = \begin{bmatrix} u_3 \\ 0 \\ u_1 \end{bmatrix}, \quad \mathbf{D}_3^i = \nabla \phi_3 = \begin{bmatrix} 0 \\ u_3 \\ u_2 \end{bmatrix}. \quad (40)$$

If the underlying factors are symmetrically distributed with unit variances and uncorrelated, then from Equation 14 we know that  $\text{Cov}(\phi_1, \phi_2) = \langle \nabla \phi_1, \nabla \phi_2 \rangle = u_2 u_3$ . The direction of most rapid change in  $\text{Cov}(\phi_1, \phi_2)$  is then

$$\nabla \text{Cov}(\phi_1, \phi_2) = \nabla \langle \nabla \phi_1, \nabla \phi_2 \rangle = \begin{bmatrix} 0 \\ u_3 \\ u_2 \end{bmatrix} = \nabla \phi_3. \quad (41)$$

Thus, the gradient of  $\phi_3$  is also the gradient of the covariance between  $\phi_1$  and  $\phi_2$ . From the symmetry of Equation 39 we can see that the same holds for any pair of these traits; the gradient of the covariance of any two traits in Equation 39 is the same as the gradient of the remaining trait. Changing any one of these traits thus leads to a consequent change in the developmental covariance of the other two. Such an association between the covariance between two traits and the value of a third trait has recently been observed in natural populations (ROFF *et al.* 2002).

### DISCUSSION

Developmental associations between two phenotypic traits arise when some of the same underlying genetic or environmental factors influence the development of both traits. In this article, I have presented a formal method for studying developmental associations on the basis of the geometry of phenotype landscapes. This approach uses tensors to capture the distribution of

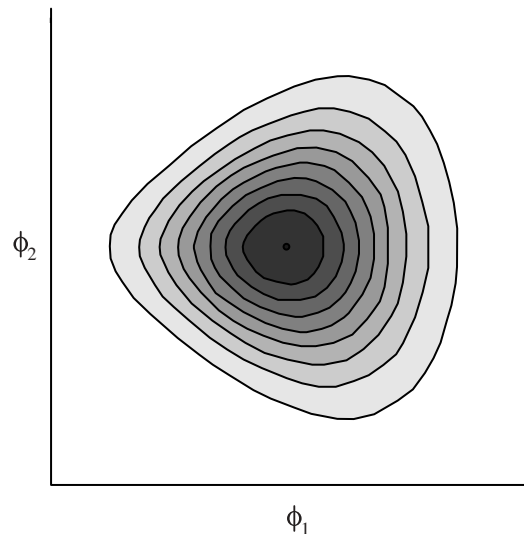


FIGURE 8.—Expected distribution of phenotypic traits for the system shown in Figure 7 given a symmetrical distribution of underlying factors with no covariance.

underlying factors (both genetic and environmental) and the interactions between these factors in development. The results are thus largely independent of how we choose to measure the underlying genetic and environmental factors.

This approach allows us to study general developmental entanglement of traits. Two traits are developmentally entangled when a developmental change that alters some moment of the distribution of one trait leads to a change in some (potentially different) moment of the distribution of the other trait. An important special case of developmental entanglement is developmental covariance between two traits, in which case the mean (first moment) of one trait is entangled with the mean of the other trait. Covariance is the only sort of developmental association possible when the underlying factors contribute additively to both traits. In the additive case, covariance between traits is determined completely by the distribution of underlying variation and the angle between the gradient vectors of the phenotype landscapes corresponding to each trait. When underlying factors interact epistatically, though, then these nonadditive interactions may contribute to phenotypic and genetic covariance.

Covariance resulting from epistatic interactions differs in at least two ways from that resulting from pleiotropy in an additive system. First, phenotypic and genetic correlation in nonadditive systems is a function of the total amount of underlying variation, because increasing underlying variation exposes more of the curvature of the phenotype landscapes. Second, epistatic interactions can produce asymmetries in the joint distribution of phenotypic traits. One result of this is that, for correlated traits, it is often easier to evolve in one direction than in the opposite direction, even if the distribution of underlying variation is symmetrical.

The effects of developmental interactions on heritability and additive genetic variance can be studied by treating the value of a character in parents as one trait and the value of the same character in their offspring as another trait and then calculating the covariance between these. When all underlying factors, genetic and environmental, contribute additively to phenotype, then heritability changes only as a consequence of changes in the distribution of underlying variation. When there are epistatic interactions, though, then heritability may evolve, either increasing or decreasing, even when there is no change in the mean value of the trait. This is relevant both to understanding phenomena like genetic assimilation and to thinking about the evolution of novel characters. If we think of a trait as simply some function of the underlying factors, and “evolutionary novelty” as the appearance of a heritable trait where none existed before, then one way that such a novel trait may arise is through a change in development that causes a previously nonheritable function of the underlying factors to become heritable.

Another consequence of epistasis is the possibility of entanglement between higher-order moments of the distributions of the traits concerned. For example, if at least one trait exhibits first-order epistatic effects, then it is possible to have a case in which stabilizing selection on one trait leads to directional change in the mean of the other trait. This may occur even when the two traits are phenotypically and genetically uncorrelated.

The idea of integration of traits resulting from shared developmental pathways has been one of the principal concepts linking development and evolution. Most authors who have discussed integration have defined it in terms of covariance between the traits involved (OLSEN and MILLER 1958; WAGNER 1990; MAGWENE 2001). Under these definitions, integration of traits would be a subset of what I am calling entanglement. The above examples show that simply saying that two traits are not integrated is not the same as saying that they are evolutionarily uncoupled. Many physiological and morphological traits are probably under strong stabilizing selection for long periods of time. Also, for some characters such as the components of articulating skeletons, selection probably acts more strongly on the ratio of different parts than on their absolute sizes. In such cases, entanglement between the variance or degree of integration of certain traits and the mean values of others may influence the direction of evolution as strongly as do genetic covariances between traits.

This article benefited greatly from comments by Bruce Walsh and two anonymous reviewers.

#### LITERATURE CITED

- BELL, A. E., and M. J. BURRIS, 1973 Simultaneous selection for two correlated traits in *Tribolium*. *Genet. Res.* **21**: 24–46.
- DE JONG, G., and A. J. VAN NOORDWIJK, 1992 Acquisition and allocation of resources: genetic (co)variances, selection, and life histories. *Am. Nat.* **139**: 749–770.
- HOULE, D., 1991 Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. *Evolution* **45**: 630–648.
- JACQUARD, A., 1983 Heritability—one word, 3 concepts. *Biometrics* **39**: 465–477.
- KEIGHTLEY, P. D., 1989 Models of quantitative variation in flux in metabolic pathways. *Genetics* **121**: 869–876.
- KOJIMA, K. I., 1959 Stable equilibria for the optimum model. *Proc. Natl. Acad. Sci. USA* **45**: 989–993.
- LANDE, R., 1979 Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. *Evolution* **33**: 402–416.
- LYNCH, M., and J. B. WALSH, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
- MAGWENE, P. M., 2001 New tools for studying integration and modularity. *Evolution* **55**: 1734–1755.
- MAYNARD SMITH, J., R. BURIAN, S. KAUFMANN, P. ALBERCH, J. CAMPBELL *et al.*, 1985 Developmental constraints and evolution. *Q. Rev. Biol.* **60**: 265–287.
- NIJHOUT, H. F., 2002 The nature of robustness in development. *BioEssays* **24**: 553–563.
- NORDSKOG, A. W., 1977 Success and failure of quantitative genetic theory in poultry, pp. 47–51 in *Proceedings of the International Conference on Quantitative Genetics*, edited by E. POLLACK, O. KEMPTHORNE and T. B. BAILY, JR. Iowa State University Press, Ames, IA.

- OLSEN, E. C., and R. L. MILLER, 1958 *Morphological Integration*. University of Chicago Press, Chicago.
- PHILLIPS, P. C., 1998 The language of gene interaction. *Genetics* **149**: 1167–1171.
- PRICE, G. R., 1970 Selection and covariance. *Nature* **227**: 520–521.
- RENDEL, J. M., 1963 Correlation between the number of scutellar and abdominal bristles in *Drosophila melanogaster*. *Genetics* **48**: 391–408.
- RICE, S. H., 1998 The evolution of development and the breaking of von Baer's laws: modeling the evolution of development with epistasis. *Evolution* **52**: 647–656.
- RICE, S. H., 2000 The evolution of developmental interactions, pp. 82–98 in *Epistasis and the Evolutionary Process*, edited by J. B. WOLF, E. D. BRODIE III and M. J. WADE. Oxford University Press, New York.
- RICE, S. H., 2002 A general population genetic theory for the evolution of developmental interactions. *Proc. Natl. Acad. Sci. USA* **99**: 15518–15523.
- RISKA, B., 1986 Some models for development, growth, and morphometric correlation. *Evolution* **40**: 1303–1311.
- ROBERTSON, A., 1966 A mathematical model of the culling process of dairy cattle. *Anim. Prod.* **8**: 95–108.
- ROFF, D. A., 1997 *Evolutionary Quantitative Genetics*. Chapman & Hall, New York.
- ROFF, D. A., S. MOSTOWY and D. J. FAIRBAIRN, 2002 The evolution of trade-offs: testing predictions on response to selection and environmental variation. *Evolution* **56**: 84–95.
- RUTLEDGE, J. J., E. J. EISEN and J. E. LEGATES, 1973 An experimental evaluation of genetic correlation. *Genetics* **75**: 709–726.
- SCHADT, E. E., S. A. MONKS, T. A. DRAKE, A. J. LUSIS, N. CHE *et al.*, 2003 Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**: 297–302.
- SIMMONDS, J. G., 1994 *A Brief on Tensor Analysis*. Springer Verlag, New York.
- WAGNER, G. P., 1990 A comparative study of morphological integration in *Apis mellifera* (Insecta, Hymenoptera). *Z. Zool. Syst. Evol.* **28**: 48–61.
- WOLF, J. B., W. A. FRANKINO, A. F. AGRAWAL, E. D. BRODIE III and A. J. MOORE, 2001 Developmental interactions and the constituents of quantitative variation. *Evolution* **55**: 232–245.

Communicating editor: J. B. WALSH

## APPENDIX

**Derivation of Equation 6:** Equation 6 is derived by noting that  $\text{Cov}(\phi_1, \phi_2) = E[\phi_1 \phi_2] - \bar{\phi}_1 \bar{\phi}_2$ . The mean value for a trait can be written

$$\bar{\phi} = \phi(\bar{u}) + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_i^k \frac{\partial^k \phi}{\partial u_{i_1} \dots \partial u_{i_k}} E[x_{i_1} \dots x_{i_k}] \quad (\text{A1})$$

(RICE 2002) and the product of two traits is

$$\begin{aligned} \phi_1 \phi_2 &= \phi_1(\bar{u}) \phi_2(\bar{u}) \\ &+ \phi_1(\bar{u}) \sum_{k=1}^{\infty} \frac{1}{k!} \sum_i^k \frac{\partial^k \phi_1}{\partial u_{i_1} \dots \partial u_{i_k}} x_{i_1} \dots x_{i_k} \\ &+ \phi_2(\bar{u}) \sum_{k=1}^{\infty} \frac{1}{k!} \sum_i^k \frac{\partial^k \phi_2}{\partial u_{i_1} \dots \partial u_{i_k}} x_{i_1} \dots x_{i_k} \\ &+ \left( \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^j \frac{\partial^j \phi_1}{\partial u_{i_1} \dots \partial u_{i_j}} x_{i_1} \dots x_{i_j} \right) \\ &\times \left( \sum_{k=1}^{\infty} \frac{1}{k!} \sum_i^k \frac{\partial^k \phi_2}{\partial u_{i_1} \dots \partial u_{i_k}} x_{i_1} \dots x_{i_k} \right). \end{aligned} \quad (\text{A2})$$

We find the expectation by multiplying Equation A2 by

the probability density function of the underlying factors and integrating over all values. The key is to note that

$$\int_{-\infty}^{\infty} \sum_i^k \frac{\partial^k \phi}{\partial u_{i_1} \dots \partial u_{i_k}} x_{i_1} \dots x_{i_k} p(x_{i_1} \dots x_{i_k}) dx_{i_1} \dots dx_{i_k} = \langle \mathbf{P}^k, \mathbf{D}^k \rangle,$$

so for  $E[\phi_1 \phi_2]$  we get

$$\begin{aligned} E[\phi_1 \phi_2] &= \phi_1(\bar{u}) \phi_2(\bar{u}) \\ &+ \phi_1(\bar{u}) \sum_{k=1}^{\infty} \frac{1}{k!} \langle \mathbf{P}^k, \mathbf{D}_1^k \rangle + \phi_2(\bar{u}) \sum_{k=1}^{\infty} \frac{1}{k!} \langle \mathbf{P}^k, \mathbf{D}_2^k \rangle \\ &+ \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{j! k!} \langle \mathbf{P}^{j+k}, \mathbf{D}_1^j \otimes \mathbf{D}_2^k \rangle. \end{aligned} \quad (\text{A3})$$

The value of  $\bar{\phi}_1 \bar{\phi}_2$  is found from Equation A1. In this case, the integral is taken before the two terms are multiplied together, so we must replace  $\mathbf{P}^{j+k}$  with  $(\mathbf{P}^j \otimes \mathbf{P}^k)$ ; both are tensors of rank  $j+k$ , but in general they will not have the same elements:

$$\begin{aligned} \bar{\phi}_1 \bar{\phi}_2 &= \phi_1(\bar{u}) \phi_2(\bar{u}) \\ &+ \phi_1(\bar{u}) \sum_{k=1}^{\infty} \frac{1}{k!} \langle \mathbf{P}^k, \mathbf{D}_1^k \rangle + \phi_2(\bar{u}) \sum_{k=1}^{\infty} \frac{1}{k!} \langle \mathbf{P}^k, \mathbf{D}_2^k \rangle \\ &+ \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{j! k!} \langle \mathbf{P}^j \otimes \mathbf{P}^k, \mathbf{D}_1^j \otimes \mathbf{D}_2^k \rangle. \end{aligned} \quad (\text{A4})$$

Subtracting A4 from A3 yields Equation 6.

**Non-Cartesian coordinates:** In this article, I have described the space of underlying factors in terms of Cartesian coordinates (*i.e.*, the coordinate axes are all at right angles to one another), in which distances are measured with a Euclidean metric [*i.e.*, the distance between points  $x$  and  $y$ , where  $x = (x_1, x_2, \dots)$  and  $y = (y_1, y_2, \dots)$ , is  $[\sum (x_i - y_i)^2]^{1/2}$ ]. Below I briefly discuss a standard result from tensor analysis (see, for example, SIMMONDS 1994) that allows us to use the results derived in this article even when our coordinate system is not Cartesian. The key is that using non-Cartesian coordinates requires that we measure the phenotype landscape and the distribution of underlying variation in different ways.

Consider, for illustration, the two-dimensional space shown in Figure A1. Any two vectors that do not point in exactly the same direction could be used as the basis for a coordinate system in this space; such vectors are called basis vectors, and we could identify any point in the space as shown in Figure A1A. The basis vectors  $\mathbf{u}_1$  and  $\mathbf{u}_2$  are referred to as the “contravariant” bases of the space. Given these vectors, we can always construct another set of basis vectors,  $\mathbf{u}_1^*$  and  $\mathbf{u}_2^*$ , that meet the following criteria:

1.  $\mathbf{u}_1^*$  is perpendicular to  $\mathbf{u}_2$  and  $\mathbf{u}_2^*$  is perpendicular to  $\mathbf{u}_1$  [*i.e.*,  $\langle \mathbf{u}_1^*, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_2^* \rangle = 0$ ].
2.  $\langle \mathbf{u}_1, \mathbf{u}_1^* \rangle = 1$  and  $\langle \mathbf{u}_2, \mathbf{u}_2^* \rangle = 1$ .

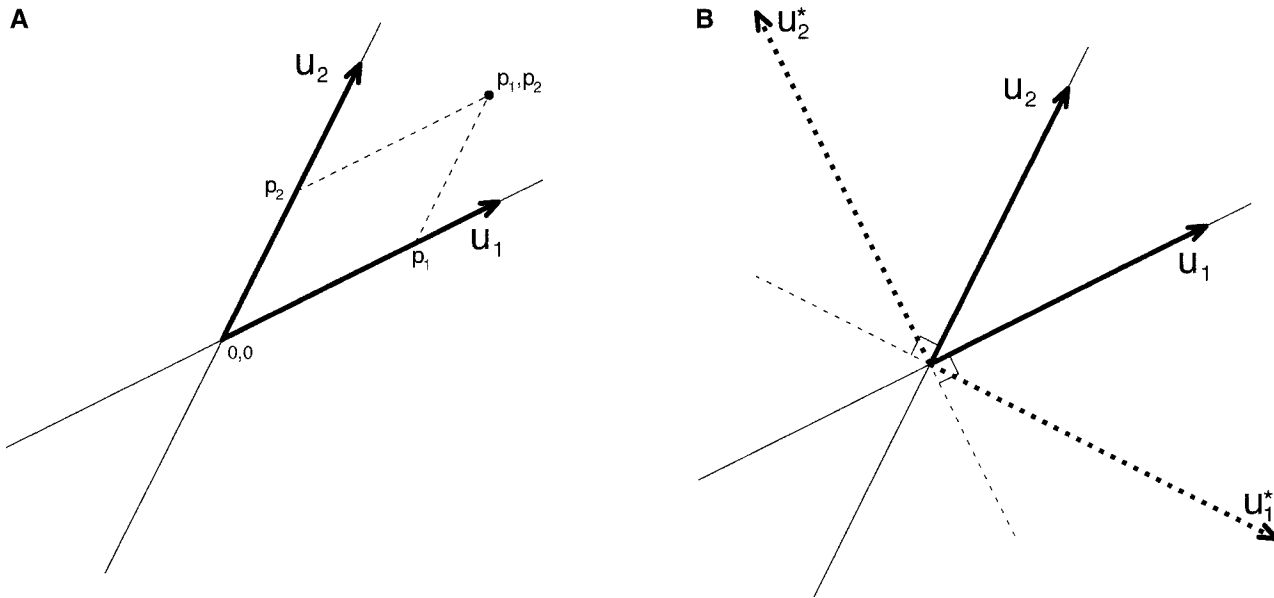


FIGURE A1.—Construction of a covariant basis from a contravariant basis. A shows how a point is located with respect to the contravariant basis vectors  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . B shows the corresponding covariant basis vectors,  $\mathbf{u}_1^*$  and  $\mathbf{u}_2^*$ .

For spaces with more than two dimensions, we extend rule 1 to require that  $\mathbf{u}_i^*$  is perpendicular to  $\mathbf{u}_j$  for all  $j \neq i$ . The vectors  $\mathbf{u}_i^*$  and  $\mathbf{u}_j^*$  are referred to as the “covariant” bases of our space. (The unfortunate terms contravariant and covariant derive from the manner in which the elements of the vectors change when we rotate the space; they have nothing whatsoever to do with the statistical concept of covariance used elsewhere in this article.) All of the equations in this article remain correct so long as we represent the phenotype landscapes (and thus measure the  $\mathbf{D}$  tensors) in terms of the contra-

variant basis vectors (whatever we choose these to be) and describe the distribution of underlying variation (the  $\mathbf{P}$  tensors) in terms of the corresponding covariant basis vectors. This requires no new information, since we can always find the covariant basis vectors from the contravariant vectors using the two rules listed above. Note that if we start out with Cartesian coordinates, in which  $\mathbf{u}_1$  and  $\mathbf{u}_2$  are at right angles and each of unit length, then these coordinates are both contravariant and covariant and we need not worry about the distinction.