Spherical Cows Grazing in Flatland:

Constraints to Selection and Adaptation

Mark Blows

School of Integrative Biology

University of Queensland

St. Lucia, Queensland 4072, Australia

m.blows@uq.edu.au

Bruce Walsh*

Department of Ecology and Evolutionary Biology

University of Arizona

Tucson, AZ USA 85721

jbwalsh@u.arizona.edu

*Corresponding Author

Version 11 Sept 2007

Header: Tales of Spherical Cows

Abstract

The vector of traits that a breeder is trying to improve and / or natural selection is acting upon has a distribution of phenotypic and breeding values that lives in a complex space. This space is not simply a sphere with equal variation in all dimensions, but rather a much more constrained structure, and these constraints have critical implications for selection responses. Here we emphasize the importance of the resurgence of interest in Fisher's geometric model of adaptation, and the necessity of taking a multivariate view of selection. We review basic matrix tools, such as the angle between two vectors, the projection of a vector into a matrix subspace, and more recent advances such as the estimation of the dimensionality of a covariance matrix, that provide different ways to quantify potential constraints on evolutionary change. A key goal of quantitative genetics is to now understand the geometry of **G** from both the point of view of mutation generating genetic variance and from selection depleting it. Initial studies using *Drosophila* have suggested that **G** may be very ill-conditioned, with a number of phenotypic dimensions displaying very little genetic variance. Such multivariate constraints may be quite important, potentially resulting in very little usable genetic variation in the direction of persistent directional selection, despite significant heritablities in each of the component traits.

Adaptation and The Multivariate Phenotype

Organisms are complex structures that are also highly variable. Couple this with the fact that scientists like to measure things such as heights, weights, bristle numbers, length of horns, size of genitalia, wing spans, and mRNA levels (to name infinitesimally few of the possibilities). Finally, throw in measures of the same character over different times and/or environments and the list of potential traits is literally endless. Thus, organisms are an essentially unmeasurable complex of traits, some of which are targets for selection while others are effectively neutral. This is the essence of the task facing evolutionary biologists when trying to understand adaptation. Even when one chooses just the smallest subset of possible traits as candidates for targets of selection, a surprisingly complex multivariate phenotype still emerges.

An additional complication is that organisms are not the mythical "spherical cow". Traits have constraints that may limit the likelihood of particular combinations of phenotypic values. Even worse, the same holds for breeding values, potentially closing off (or at least rendering extremely inefficient) certain paths for improved fitness. Thus, any view of adaptation has a strong geometric component. This view dates back to D'Arcy Thompson (1917) and the founders of the modern synthesis, R. A. Fisher (1930), Sewall Wright (1932), and G. G. Simpson (1944). In the modern analysis of adaptation, we also take a geometric view, considering the distribution of values in both phenotypic and genetic (breeding value) space. Here we examine some recent theoretical developments in this geometric view (both genetical and statistical), and some recent empirical work that suggests a central role for the geometric representation of constraints to selection and adaptation.

Spherical Cows: Geometric Models for the Adapativeness of New Mutations

Geometric descriptions of organismal morphology began with the classic treatment of Thompson (1917), wherein different trait values can be thought of as axes in some morphological space. We can superimpose a fitness surface (Wright 1932, Simpson 1944, Lande 1979) upon this space, adding an axis that represents the mean fitness of any particular combination of trait values. On such a fitness space, Fisher (1930) suggested that the actual number of independent traits under selection has important consequences for adaptation. His model has subsequently been investigated by a number of other workers (e.g., Kimura 1983, Leigh 1987, Rice 1990, Hartl and Tabes 1996, Orr 1998, 2000). Fisher's model, in a classic spherical cow approximation, assumes an overly simplistic geometry for the resulting fitness surface, with a single optimal fitness peak (at the optimal point θ) with fitness uniformly decreasing as we increase the Euclidean distance from θ (Figure 1). Fisher's interest was how likely a new mutation was to be adaptive (increasing its fitness relative to the current value). In particular, given a phenotype at point z, which is at distance d from the optimal value θ, what is the probability that a new mutation with effect r (the distance between the wild-type and mutant phenotypes) increases fitness? For a two-dimensional fitness surface, Fisher's model gives the contour of equal fitness of the current phenotype as a circle, allowing a simple geometric argument to be used to determine how likely it is that a mutation ends up with higher fitness (Figure 1). Fisher's key insight was that, for phenotypes far away from the optimum value, large moves can have a higher probablity of adapation than small moves. However, as we approach the optimum, large moves almost always have reduced fitness. In particular, it is the scaled ratio r/d, the fraction of the move relative to the distance from the optimum, that characterizes the probablity of adaptation.

— Figure 1 here –

For n traits, Fisher assumed that the contours of equal fitness are hyper-spheres around the

optimum value, in which case the probability of adaptation is determined by the single quality

$$x = \frac{r\sqrt{n}}{2d} \tag{1}$$

which now scales the relative move r/d by the dimensionality n of the fitness space. For large n, Fisher (1930) showed that the probability of such a mutation being favorable (increasing fitness) is

$$p_{fav} = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-y^2/2)dy = 1 - erf(x) \tag{2}$$

where erf is the error function (Rice 1990 gives exact expressions for any n). Thus, as noted by Fisher, increasing the number of independent dimensions of selection (increasing n) makes a new mutation increasingly unlikely to be adaptive. The direct (but often unstated) implication of this result is that organisms must be somewhat constrained in the number of independent dimensions that fitness can act upon.

Kimura (1983) and Orr (1998, 2000) develope an important extension of Fisher's model. Fisher simply considered the probability of a favorable mutant. However, if a favorable new mutation is introduced as a single copy in an infinite population, then its probability of fixation is just $2s$, twice its selective advantage as a heterozygote, so that fixation of most favorable new mutations is unlikely. Thus, while mutations of larger effect (x larger) are increasingly unlikely to be favorable, when they are, they may have a significantly larger effect, and hence a higher probability of fixation. The quantity of interest is the fixation rate, the product of the average selective advantage (when favorable) times the probability of being favorable (Equation 2). For example, as the mutational effect r becomes infinitesimally small, the probability of it being advantageous approaches 50%, but its average advantage (when favorable) also approaches 0 (as $s \to 0$). Orr (2000) showed that the optimal "size" for a mutation (to maximize this product) is roughly $x \simeq 0.925$, so that during an adaptive walk trying to move from z to θ through a successive series of mutations, the optimal mutation size gets smaller with each successive fixation (i.e., as d gets closer to zero). Specifically, from Equation 1, the optimal mutation size is approximately

$$r_{opt} \simeq \frac{0.925 \cdot 2 \cdot d}{\sqrt{n}} \simeq 1.85 \cdot \frac{d}{\sqrt{n}}$$

Orr further showed that there is a considerable cost to complexity (dimensions of selection n), with the rate of adaptation (favorable mutation rate times fixation probability) declining significantly faster that $1/n$. Thus, the constraint on dimensionality may be much more severe than originally suggested by Fisher.

Much evolutionary angst has been generated over the "cost of complexity" given these results. However, Fisher's model makes the simplifying assumption that all traits have the same fitness curvature, which is highly unlikely. More likely is a fitness surface with a few dimensions under strong selection (large curvature) and a very large number of dimensions under very weak selection (low curvature) (Barton 1990; Johnson and Barton 2005). Rice (1990) has examined such cases where the curvature of the fitness surface is not constant across all dimensions. The probability of adaptation on these surfaces depends upon their "effective curvature", roughly the harmonic mean of the individual curvatures. Recalling that the harmonic mean is dominated by small values, it follows that the probability of adaptation is likewise dominated by those fitness surfaces with low curvature (weak selection). However, on such surfaces, s is small, and hence the fixation probability small. Clearly, adaptive walks in such variable-curvatures surfaces is an interested area for much further work. The key, however, is that while it is sophisticated to think about the geometric space generated by a fitness surface, we may obtain very misleading results if our assumed geometry is oversimplified.

Multivariate Phenotypes and Selection Response

As we have stressed, all selection is multivariate. Even in those (rare) cases where artificial selection is focused on just a single trait, natural selection on other (potentially) correlated traits is occurring as well, and this additional selection can help or (more likely) hinder the artificial selection objectives. In natural populations, wherein organisms are attempting to adapt to ever-changing environments, the adaptive target may be constantly shifting. Indeed, as organisms evolve, they necessarily change the environment around them, and hence organisms must evolve just to keep pace with other organisms with which they compete for resources. This is van Valen's (1973) red-queen hypothesis, wherein (after Lewis Carroll's Red Queen in *Through the Looking Glass*) one must run just to stay in place.

Attempts to model multivariate selection response have two components. The first, prediction of response given the nature of selection, is a concern shared by both breeders and evolutionary biologists. For example, Smith (1936), Hazel (1943), and others have developed selection indices to maximize some linear combination of traits given a desired set of weights. Evolutionary biologists face a second task, estimating the actual nature of selection. Even in an apparently simple univariate setting, this approach is fraught with peril. Particularly distressing is that there are numerous examples of well-studied natural populations (most monitored for over a decade) that show significant heritability for a trait, a (relatively) constant selection differential on that trait, and yet while h^2 and S are both significant, the simple breeders equation ($R = h^2 S$) is not followed, as no response is seen (Merilä et al 2001). One (but not the only) explanation is that a third trait, influencing both fitness and our focal trait, can generate a focal trait-fitness correlation. If the causative trait is not heritable, there is no response. A classic example is antler size in Red Deer, where males with larger antlers gather more mates, and antler size is highly heritable. However, analysis using animal model BLUP shows that there is no direct selection in breeding values for antler size (Kruuk et al. 2002). Rather, the suggestion is that an animal's nutritional state, if high, allows them to grow larger antlers. Further, if nutritional state is high, males are also better fighters, getting more mates. This creates an antler-fitness correlation entirely due to the effect of a third trait, nutritional state, on both.

A formal theory for multivariate selection was proposed by Lande (1979), ironically at roughly the same time that multivariate BLUP was proposed (Henderson and Quaas 1976). The multivariate version of the breeders' equation relates the vector $\mathbf{R}$ of responses and the vector $\mathbf{S}$ of selection differentials through the phenotype $\mathbf{P}$ and genetic (breeding value) $\mathbf{G}$ covariance matrices,

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} \tag{3}$$

Note that in the multivariate version $\mathbf{G}\mathbf{P}^{-1}$ replaces $h^2 = \sigma_A^2(1/\sigma_z^2)$. The multivariate breeders' equation follows under the assumption that the joint distribution of phenotypes and breeding values is multivariate normal.

Robertson (1966) and Price (1970) showed that the (univariate) selection differential is mathematically equivalent to the covariance between trait value and fitness. Thus, $\mathbf{S}$ is a vector of covariances, the i-th element being the covariance between trait i and fitness, $S_i = \sigma(z_i, w)$, and from the theory of multiple regression it follows that

$$\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S} \tag{4a}$$

is the vector of weights of the multiple regression of fitness w on the vector $(z_1, z_2, \cdots, z_n)$ of phenotypes,

$$w = a + \sum_{i=1}^{n} \beta_i z_i + e_i \tag{4b}$$

β_i is called the directional selection gradient on the ith trait, and is a measure of the amount of direct selection acting on that trait, after (phenotypic) correlations with all of the other traits in the

sample have been removed (Pearson 1903, Lande and Arnold 1983). An important caveat is that unmeasured traits under direct selection that are phenotypically correlated with trait i can create a false picture of the nature of selection on i.

Putting all of this together, we have the Lande Equation, relating a measure of phenotypic selection $\boldsymbol{\beta}$ (removing the effects of phenotypic correlations) with the usable space of breeding values (given by $\mathbf{G}$),

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} \tag{5}$$

If phenotypes are multivariate-normally distributed, then Lande (1979) shows that the direction the current mean must move to maximize the local change in mean population fitness is given by the vector $\boldsymbol{\beta}$. Formally, $\boldsymbol{\beta}$ is a gradient with respect to (log) mean population fitness as a function of the trait means. Thus fitness change is maximized by moving the population in the direction $\boldsymbol{\beta}$, but genetics *constrains* this response through the variance-covariance matrix $\mathbf{G}$. When a vector is multiplied by a matrix, it results in a rotation and scaling. Any rotation results in the direction of change not being along the optimal direction $\boldsymbol{\beta}$ to increase fitness, while scaling changes the relative rate of response along the final direction. This is also a well-known lesson for index selection: specification of the breeding objective (the vector of economic weights on a selection index) does not fully specific the direction of genetic change.

Equations 4a and 5 show how constraints arise in multivariate selection and how (under multivariate normality) covariance matrices quantify these constraints. The first potential constraint is on phenotypes. Suppose selection is directly acting on k traits, but the resulting phenotypic covariance matrix $\mathbf{P}$ for these traits is singular (or nearly so). In such cases, the effects of phenotypic selection will not fully translate into direct selection on individual traits. If $\mathbf{P}$ is ill-conditioned (nearly singular), then small random fluctuations may result in large changes in individual β values, decreasing the efficiency of selection response. When $\mathbf{P}$ is singular, then so is $\mathbf{G}$ (Pease and Bull 1988), as $\mathbf{P} = \mathbf{G} + \mathbf{E}$, the sum of the genetic and environmental covariance matrices. However, the converse is not true, and $\mathbf{P}$ can be of full rank even when $\mathbf{G}$ is singular. Indeed, most discussion of constraints in the response to selection focus on constraints imposed by $\mathbf{G}$, but this may be only part of the story. Biologically, singularities in a covariance matrix can arise when a trait has essentially no (phenotypic or genetic) variation and / or (more likely) when traits highly covary with each other.

Living in Flatland: The Misleading Univariate World

Selection and response are intrinsically high-dimensional phenomena. However, the focus of breeders, and even evolutionary biologists, is on a much restricted subset of this space, often one or just a few traits. The problems of mapping from a high dimensional space into a lower dimensional one were first noted in Abbott Abbott's (1884) wonderful *Flatland*. In Abbott's world, geometric shapes (such as triangles) living in two dimensions take on different appearances to their fellow Flatlandians as their orientation changes. Visitors from 3-D space sojourning through Flatland have even more dramatic changes, starting out as a circle of increasing size, which reaches maximal size and then shrinks back to zero as the sphere passes through its 2-D projection in Flatland. As biologists (and breeders) we tend to work in flatland, ignoring (indeed, we are almost always ignorant of) the full geometry underlying the traits we are intending to follow and / or modify, and instead work with a projection of the full space onto a lower dimensional one.

As early as Dickerson (1955), some of the consequences of a univariate view of multivariate selection were noted. Dickerson (in the context of a selection index) remarked that each trait in an index may have a significant heritability, but the resulting heritability of the index itself may be zero (or nearly so). Thus, with multivariate selection there can be genetic variation (i.e., non-zero heritabilities) in all traits under selection, but little or no genetic variation along direction that selection is trying to move the population. To see this point, consider the following simple example. Suppose the additive variance of traits 1 and 2 are 20 and 40, respectively. Further, suppose that

there is direct selection on both traits, with selection favoring a two unit change in trait 1 for each one-unit decrease in trait 2, e.g.,

$$\boldsymbol{\beta} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

If we consider each trait separately, in a fully univariate framework ignoring all other traits, the selection response for trait i is just $R_i = \sigma_A^2(i)\beta_i$, which would be $10 \cdot 2 = 20$ for trait one and $40 \cdot (-1) = -40$ for trait two. When $\sigma(A_1, A_2) = 0$, these are indeed the responses predicted from the multivariate breeders' equation,

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \begin{pmatrix} 10 & 0 \\ 0 & 40 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 20 \\ -40 \end{pmatrix}$$

While there is response in both traits, there is actually a (minor) constraint here (as $\mathbf{R}$ does not point in the same direction as $\boldsymbol{\beta}$), a point we return to shortly.

Now suppose that the genetic covariance between traits 1 and 2 is $\sigma(A_1, A_2) = 20$, giving

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \begin{pmatrix} 10 & 20 \\ 20 & 40 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Thus, even though there is considerable additive genetic variation in both traits, there is no response. This occurs because $\mathbf{G}$ as a zero eigenvalue, whose associated eigenvector exactly corresponds to our $\boldsymbol{\beta}$. There is no additive genetic variance along this particular direction, and hence no response. This is an example of an absolute constraint (Pease and Bull 1988). Quantitative constraints can arise when $\boldsymbol{\beta}$ is only a few degrees away from an eigenvector associated with a zero eigenvalue, in which case the resulting response would be very small. In theory, especially in a rapidly changing environment, a quantitative constraint could result in population extinction, as although the population can response in the desired direction, it may not be able to do so quickly enough to keep pace with environmental change.

One straightforward way to quantity the constraints on the response imposed by $\mathbf{G}$ is to consider the angle θ between the optimal $\boldsymbol{\beta}$ and actual response $\mathbf{R}$. Recall that the angle θ between two vector is given by

$$\cos(\theta) = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||\ ||\mathbf{y}||}$$

Hence the constraint angle (between the desired direction $\boldsymbol{\beta}$ and the actual direction $\mathbf{R}$) is

$$\theta = \cos^{-1}\left(\frac{\mathbf{R}^T\boldsymbol{\beta}}{||\mathbf{R}||\ ||\boldsymbol{\beta}||}\right) \tag{6}$$

There is no constraint if $\theta = 0$ ($\mathbf{R}$ and $\boldsymbol{\beta}$ have the same orientation), while $\theta = 90$ degrees corresponds to an absolute constraint. Now let's reconsider the response above with no genetic covariance between the traits. Selection ($\boldsymbol{\beta}$) is maximized by a two unit response in trait one for every (negative) unit of response in trait two. However, the actual response in this case was in a (slightly) different direction, doubling the response in trait two for every unit of response in trait one (Figure 2). Equation 6 gives $\theta = \cos^{-1}(0.8)$, for θ = 37 degrees. Thus, the actual response is rotated 37 degrees from the optimal direction for selection response by $\mathbf{G}$. Notice that this (mild) constraint occurs even though the two traits are genetically uncorrelated. Trait two has four times the genetic variance of trait one, and hence gives a larger response per unit of selection than trait one. Thus, even a diagonal (and non-singular) $\mathbf{G}$ matrix can impose constraints to response, although these are expected to be mild unless the largest values of β_i are associated with the smallest G_{ii}.

— Figure 2 Here —

Constraints and Consequences

While the inherent genetic constraints imposed by $\mathbf{G}$ have long been appreciated by evolutionary biologists (e.g., Lande 1979, Pease and Bull 1988), the general view of many was that the conditions for $\boldsymbol{\beta}$ to align with a zero-eigenvalue were rather pathologic and hence unlikely (by chance) to happen. However, additive genetic variation tends to erode under constant directional selection. Hence, if a trait has been under long-term selection to move in a particular direction, it is possible that any (initially) usable genetic variation along this direction has been largely eroded, with any residual variation likely generated since selection started (such as by mutation and / or recombination). Thus eigenvectors with near-zero eigenvalues along directions of response may be an evolved phenomena from persistence selection, rather than rare chance events.

How common are genetic constraints? Unfortunately, there are not a sufficient number of studies to offer any comprehensive view, for two reasons. First, if one is focusing on two or three traits, and yet the actual constraint is in a higher dimensional space, we might not observe any restrictions on the lower dimensional projection we are considering (Pease and Bull 1988). Second, as we have seen above it is important to know both the genetic basis of the traits under consideration and the direction of selection, and there are not many studies that estimate both $\boldsymbol{\beta}$ and $\mathbf{G}$ for a vector of traits under natural selection.

Recently, Blows et al. (2004) examined 8 cuticular hydrocarbons (CHC) in the Australian fruit-fly *Drosophila serrata*, which are important cues in mate choice. Both $\mathbf{G}$ for these traits, as well as their associated gradients $\boldsymbol{\beta}$ under sexual selection (mate choice), were estimated. In a large half-sib design, $\mathbf{G}$ was found to be ill-conditioned, with its first two eigenvalues accounting for 78 percent of the total additive genetic variation. The resulting two leading eigenvectors ($\mathbf{g}_1$ and $\mathbf{g}_2$) of $\mathbf{G}$ and $\boldsymbol{\beta}$ were as follows:

$$\mathbf{g}_1 = \begin{pmatrix} 0.232 \\ 0.132 \\ 0.255 \\ 0.536 \\ 0.449 \\ 0.363 \\ 0.430 \\ 0.239 \end{pmatrix}, \qquad \mathbf{g}_2 = \begin{pmatrix} 0.319 \\ 0.182 \\ 0.213 \\ -0.436 \\ 0.642 \\ -0.362 \\ -0.014 \\ -0.293 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} -0.099 \\ -0.055 \\ 0.133 \\ -0.186 \\ -0.133 \\ 0.779 \\ 0.306 \\ -0.465 \end{pmatrix}$$

As (one) measure of constraints on the evolution of this vector of traits, consider the angle θ between the direction of maximal genetic variation ($\mathbf{g}_1$) and the optimal direction favored by selection ($\boldsymbol{\beta}$). From Equation 6,

$$\theta = \cos^{-1}\left(\frac{\mathbf{g}_1^T\boldsymbol{\beta}}{||\mathbf{g}_1||\ ||\boldsymbol{\beta}||}\right) = \cos^{-1}\left(\frac{0.147496}{\sqrt{0.99896 \cdot 0.999502}}\right) = \cos^{-1}(0.1476)$$

Giving $\theta = 81.5$ degrees. Thus, the vector of maximal genetic variation and the vector of optimal response are almost at right angles. Likewise, the angle between $\mathbf{g}_2$ and $\boldsymbol{\beta}$ is $\theta = 99.65$ degrees. For this population, there is very little standing additive genetic variation in the direction of the optimal selection response. While all of the CHC traits showed significant genetic variation, and indeed have responded to selection under a number of environmental conditions in different experiments, there is very little usable genetic variation in the direction that sexual selection is trying to move the population.

Assuming $\mathbf{G}$ remains (relatively) constant, can we relate population divergence to any feature of $\mathbf{G}$? Schluter (1996) suggested that we can, as he observed that populations tend to divergence along what he called genetic lines of least resistance, $\mathbf{g}_{max}$ (the first principal component of $\mathbf{G}$). Schluter looked at morphological divergence data in a small set of vertebrates (stickleback fish, mice, and three species of birds), and observed that populations tend to divergence in the direction of $\mathbf{g}_{max}$, specifically the angle between the vector of between-population divergence in means and $\mathbf{g}_{max}$ was small.

Thus, there is at least some empirical evidence that populations tend to evolve along lines of least genetic resistance (i.e., lines of maximal genetic variance). There are two ways to interpret this observation. The first is that such lines constrain selection, with departures away from such directions being difficult. The second is that such lines are also the directions on which maximal genetic drift is expected to occur (as the between-mean variance is proportion to the total amount of genetic variation). To see this last point, consider a trait entirely under random genetic drift. Using a Brownian motion model as an approximation (Lande 1976, 1979), the expected vector of divergences in the vector of traits from their starting value is multivariate-normally distributed, so that if the population starts at mean vector $\boldsymbol{\mu}$ in generation 0, the distribution in the vector of means at generation t is

$$\boldsymbol{\mu}(t) \sim MVN\left(\boldsymbol{\mu}, \frac{t}{2N_e} \cdot \mathbf{G}\right) \tag{7}$$

Maximal divergence is along the lines of greatest variation in $\mathbf{G}$, i.e., along the first few eigenvectors of $\mathbf{G}$.

Both of these explanations involve drift, but their distinction is a lack of adaptive response (despite pressure for such a change) versus random drift in phenotypic space. An interesting perspective on this problem is offered from considerations of the divergence in two species of Australian rainbow fish (genus *Melanotaenia*) that each have populations differentially adapted to lake vs. stream hydrodynamic environments (McGuigan et al. 2005). Several morphological landmarks for hydrodynamic morphology were followed as a vector of traits. Divergence between species, as well as divergence among replicate hydrodynamic populations within each species, followed Schluter's results (small angular departures from the vector $\mathbf{d}$ of divergent means and $\mathbf{g}_{max}$). However, hydrodynamic divergence between lake versus stream populations within each species were along directions that were quite removed from $\mathbf{g}_{max}$ (as well as the other eigenvectors of $\mathbf{G}$ that described most of the genetic variation). Thus, the between- and within-species divergence within the same hydrodynamic environment are consistent with drift, while hydrodynamic divergence within each species had to occur against a gradient of very little genetic variation. Of course, one cannot rule out that the adaptation to these environments resulted in a depletion of genetic variation along these directions. Indeed, this may indeed be the case, and is a point we return to below.

It is worth noting here that empirical investigations have yet to incorporate further elaborations that may help determine how $\mathbf{G}$ may influence population divergence. The first of these involves taking full advantage of the partitioning of the additive genetic variance within and among populations. Prout and Barker (1993, Spitze 1993) developed the approach of comparing the level of among-population genetic variance in quantitative traits (Q_{st}) with genetic variation measured by netural markers (F_{st}) to test for a signal of population divergence as a consequence of selection. This approach has the potential to be expanded to the estimation of the among population $\mathbf{G}$ matrix, that in turn should enable a much closer examination of how the within-population $\mathbf{G}$ matrix influences population divergence. The second involves the issue of variation among populations in selection. Although among-population variation in selection has received little empirical attention (Arnold et al 2001), Zeng (1988) has developed a framework for determining how such variation will influence population divergence that relies upon the estimation of among-population variance-covariance matrices for selection gradients. Clearly, we are just at the beginning of investigations of how multiple traits diverge among populations.

Using Matrix Subspace Projection to Measure Constraints

Schluter's approach only considers the leading eigenvector $\mathbf{e}_1$ of $\mathbf{G}$. If its associated eigenvalue λ_1 dominates all of the others (and hence accounts for most of the variance), this single vector may be an adequate description of the space of usable genetic variation. More generally, we might expect the first few eigenvalues together may account for most of the variation, so that focusing only on the largest may miss a significant fraction of the variation. In such cases, matrix subspace projection provides a more realistic description of the usable subspace contained within $\mathbf{G}$ (Blows et al. 2004).

A second advantage of using a subspace projection deals with the common problem that the $\mathbf{G}$ matrix is often ill-conditioned, in that $\lambda_{max}/\lambda_{min}$ (the ratio of the largest to smallest eigenvalues) is large. In such cases (as well as others!) estimation of the $\mathbf{G}$ matrix may result in estimates of eigenvalues that are very close to zero or even negative (Hill and Thompson 1978). Negative estimates arise due to sampling (Hill and Thompson 1978), but values near zero may reflect the true biology in that there is very little variation in certain dimensions. One can extract (estimate) a subspace of $\mathbf{G}$ that accounts for the vast majority of useable genetic variation by, for example, taking the leading k eigenvectors. It is often the case that $\mathbf{G}$ contains several eigenvalues whose associated eigenvectors account for almost no variation (i.e, $\lambda_i/\mathrm{tr}(\mathbf{G}) \simeq 0$, recalling that the trace tr of a matrix [the sum of its diagonal elements] equals the sum of its eigenvalues, $\mathrm{tr}(\mathbf{G}) = \sum_i G_{ii} = \sum_i \lambda_i$). In such cases, most of the genetic variation resides on a lower-dimensional subspace.

Quantifying genetic constraints within this subspace follows using the projection of the full space of $\mathbf{G}$ into this subspace (this is just the matrix extension to the projection of one vector onto another). Suppose we have included the first k eigenvectors in our analysis. These form a projection matrix by first defining the matrix $\mathbf{A}$, where

$$\mathbf{A} = (\,\mathbf{g}_1, \quad \mathbf{g}_2, \quad \cdots, \mathbf{g}_k\,) \tag{8}$$

so that the $\mathbf{A}$ matrix consists of the first k eigenvectors of $\mathbf{G}$. The resulting projection matrix becomes

$$\mathbf{P}_{roj} = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T \tag{9a}$$

and in particular, the projection of the optimal vector of selection response $\boldsymbol{\beta}$ onto this subspace of $\mathbf{G}$ (the subspace that essentially contains all of the useable additive variation) is given by

$$\mathbf{p} = \mathbf{P}_{roj}\boldsymbol{\beta} = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\boldsymbol{\beta} \tag{9b}$$

Let's reconsider the *Drosophia serrata* data on the eight CHC traits involved in mate choice. The first two eigenvalues account for roughly 80% of the total variation in $\mathbf{G}$, i.e., $(\lambda_1 + \lambda_2)/\sum\lambda_i = 0.78$. The resulting $\mathbf{A}$ matrix becomes

$$\mathbf{A} = (\,\mathbf{e}_1, \mathbf{e}_2\,) = \begin{pmatrix} 0.232 & 0.319 \\ 0.132 & 0.182 \\ 0.255 & 0.213 \\ 0.536 & -0.436 \\ 0.449 & 0.642 \\ 0.363 & -0.362 \\ 0.430 & -0.014 \\ 0.239 & -0.293 \end{pmatrix}$$

Applying Equation 9a gives an 8×8 projection matrix (not show here), and Equation 9b gives the

projection vector $\mathbf{b}$ of $\boldsymbol{\beta}$ onto the subspace given by $\mathbf{A}$ as

$$\mathbf{p} = \mathbf{P}_{roj}\boldsymbol{\beta} = \begin{pmatrix} -0.0192 \\ -0.0110 \\ 0.0019 \\ 0.1522 \\ -0.0413 \\ 0.1142 \\ 0.0658 \\ 0.0844 \end{pmatrix}$$

The angle θ between $\boldsymbol{\beta}$ and the projection $\mathbf{p}$ of $\boldsymbol{\beta}$ into the subspace of the genetic variance is given by

$$\theta = \cos^{-1}\left(\frac{\mathbf{p}^T\boldsymbol{\beta}}{||\mathbf{p}||\,||\boldsymbol{\beta}||}\right) = \cos^{-1}(0.223) = 77.1 \text{ degrees}$$

Thus the direction of optimal response is 77 degrees away from the genetic variation described by this subspace (which spans 78% of the total variance).

A second recent study examined a vector of CHC traits (again involved in mate choice) in *Drosophila bunnanda*, a recently described northeastern Australian sibling species of *D. serrata* (Van Homright et al. 2007). A total of 15 potential candidate CHC were used, and in mating trails, 9 of these appear to be involved in mate choice. Further, an estimate of the amounts of selection on these nine traits ($\boldsymbol{\beta}$) can be directly estimated from these mate choice experiments. The resulting estimated $\mathbf{G}$ for these traits had 98% of the total genetic variation in the first five principal components (the first four had 95% of the total variance). The angle between $\boldsymbol{\beta}$ and its projection into the 5-dimensional subspace of $\mathbf{G}$ was 88.2 degrees. Likewise, if the first four PCs were considered for the subspace, the projection is even more constrained, being 89.1 degrees away for $\boldsymbol{\beta}$. Hence, in this second example, the constraints were even stronger. When the entire space of $\mathbf{G}$ is considered, the resulting angle θ between $\mathbf{R}$ and $\boldsymbol{\beta}$ is 67 degrees.

The Dimensionality of the G Matrix

An obvious question relating to constraints is the actual dimensionality of the genetic covariance matrix $\mathbf{G}$. If the rank of $\mathbf{G}$ is much less than the number of measured traits n, then there are considerable absolute constraints. Further, even if $\mathbf{G}$ is of full rank, if several of the eigenvalues are near zero, this also imposes a significant constraint on some paths of response. Recalling that a singular phenotypic covariance matrix $\mathbf{P}$ also imposes constraints on adaptation, comparison of the ranks of both $\mathbf{P}$ and $\mathbf{G}$ should formally be done to examine the degree to which the constraints are genetic (McGuigan and Blows 2007). For example, if both $\mathbf{P}$ and $\mathbf{G}$ are singular, but of similar rank, then genetic constraints may simply be reflections of the phenotypic constraints. In contrast, if $\mathbf{P}$ is of full rank, but $\mathbf{G}$ is not, then genetic constrains are important. For example, McGuigan and Blows (2007) examined ten wing traits in *Drosophila bunnanda*. The phenotypic space had ten full supported dimensions, while the supported dimensionality of $\mathbf{G}$ for this sexually-dimorphic trait was five in males and two in females. Thus, there appears to be significant genetic constraints on these traits that are not the result of a constrained phenotypic space.

For obvious reasons of power, accurately estimating the dimension of $\mathbf{G}$ is not trivial. For example, even if the true matrix is of full rank, it is not unusual (especially if the dimensionality is large) for random samples drawn from this $\mathbf{G}$ to have not just zero estimates of eigenvalues, but negative ones (Hill and Thompson 1978, Amemiya 1985). Thus it can be very difficult to estimate the actual number of positive eigenvalues of the true $\mathbf{G}$ given a sample covariance matrix. Three different approaches to estimating dimensionality have been applied to genetic covariance matrices.

First, Amemiya (1985) develops a procedure based on estimating the characteristic roots λ_i of the equation

$$\det(\mathbf{M}_b - \lambda \mathbf{M}_w) = 0 \tag{10}$$

where $\mathbf{M}_b$ and $\mathbf{M}_w$ are the between- and within-mean square matrices. If all of the λ_i roots for Equation 10 are ≥ 1, then our estimate $\mathbf{M}_b$ of the genetic covariance matrix will be non-negative definite, otherwise, it will be indefinite. Thus, Amemiya's procedure successively tests for how many of the λ_i are indeed significantly greater (or equal) to one. If the first k are, but the $k+1$st is not, then the estimate of the dimensionality of $\mathbf{G}$ is given by k. When this approach is applied to the eight *Drosophila serrata* CHC traits mentioned previously, the dimensionality of $\mathbf{G}$ is not significantly greater than two (Hine and Blows 2006). Likewise, the dimensionality of $\mathbf{G}$ for the nine *D. bunnanda* CHCs was examined by Van Homright et al. (2007), finding strong support for 4 dimensions, and marginal support ($p = 0.081$) for a fifth.

A second approach, based on bootstrap resampling (at the sire family level for a half-sib design) was used by Mezey and Houle (2005), who examined 20 *Drosophila melanogsater* wing traits. These authors found that the resulting $\mathbf{G}$ matrix could not be shown to be significantly less than full rank. Finally, Kirkpatrick and Meyer (2004) and Meyer and Kirkpatrick (2005) have suggested estimating the principal components (PCs) directly and in successive fashion. In this factor-analytic approach, the dimensionality of $\mathbf{G}$ is the number of PCs that are added until the next does not significantly improve model fit. There are several powerful features of this approach. First, the resulting covariance matrix is guaranteed to be positive-definite (by construction). Further, the actual number of parameters that need to be estimated can be greatly reduced (on the order $k \cdot n$ for k PCs versus roughly order n^2 for estimation of the full matrix.

Preliminary simulations studies (Hine and Blows 2006) find that the bootstrap method almost always overestimates the true dimensionality, while Amemiya's method and factor-analytic modeling do a reasonable job of estimation when heritability is high. When traits have low heritability, Amemiya's method has low power, resulting in underestimation of the true dimensionality.

Evolution Under Constraints or Evolution of Constraints?

$\mathbf{G}$ both constrains selection and also evolves under selection. Over short time scales, if most alleles have modest effects, $\mathbf{G}$ changes due to selection generating linkage disequilibrium. It is important to note that under the infinitesimal model (and hence no change in allele frequencies), selection still generates considerable disequilibrium, even among unlinked loci (Bulmer 1971). Indeed, any selection (such as stabilizing or any form of directional selection) that generates a decrease in the variance following selection generates negative disequilibrium, meaning that gametes containing favorable alleles are less frequent than expected simply from the product of their allele frequencies. In particular, assuming only directional selection (no curvature in the fitness surface), the within-generation change in the additive-genetic covariance matrix under the infinitesimal model is given by

$$\Delta\mathbf{G} = -\mathbf{G}\boldsymbol{\beta}\boldsymbol{\beta}^T\mathbf{G} = -\mathbf{R}\mathbf{R}^T \tag{11a}$$

(Tallis and Leppard 1988, Lande 1979, Phillips and Arnold 1989). Thus, the (within-generation) change in the additive genetic covariance between traits i and j is

$$\Delta G_{ij} = \Delta\sigma(A_i, A_j) = -R_i R_j \tag{11b}$$

Equation 11b shows that election decreases the additive genetic variance at traits responding to selection (Bulmer 1971) and decreases the covariances if traits i and j respond in the same direction, while increasing the covariance if i and j respond in different directions. The net result is that linkage disequilibrium increases any initial constraints. A simple way to see this is to consider selection

on the index given by Equation 4b, $I = \sum_i z_i \beta_i$. Selection on this index (which is the predicted fitness) results in deceased additive variance in this composite trait (Bulmer 1971). Thus, as pointed out by Shaw et al. (1995), if one estimates $\mathbf{G}$ by first having several generations of random mating in the laboratory under little selection, existing linkage disequilibrium decays, and the resulting estimated $\mathbf{G}$ matrix may show less of a constraint than the actual $\mathbf{G}$ operating in nature (with its inherent linkage disequilibrium).

The second avenue for change in $\mathbf{G}$, allele frequency change, occurs over longer time scales (unless there are major genes). Unlike the relatively simply dynamics under linkage diseequilibrium (Equation 11), effects on $\mathbf{G}$ from allele frequency change are extremely unpredictable (Bohren et al. 1966). A general qualitative trend was suggested by Lush (1948) and Lerner (1950). They predicted that selecting two traits in the same direction results in a negative change in their genetic covariance as alleles with desirable $(++)$ and undesirable $(--)$ pleiotropic effects on the two traits are fixed (and lost) by selection leaving only alleles segregating $+-$ pleiotropic effects. While compelling, there are counterexamples to this simple prediction, e.g., Sheridan and Barker (1974). Thus, during the course of selection, $\mathbf{G}$ itself changes, often so as to increase constraints (as usable variation continues to be removed).

Hence, it is certainly not surprising that little usable genetic variation may remain along a direction of persistence directional selection. What is surprising, however, is that considerable genetic variation may exist along other directions. The quandary is not why is there so little usable variation but rather why is their so much. Indeed, quantitative genetics is in the embarrassing position as a field of having no models that adequately explain one of its central observations – genetic variation (measured by single-trait heritabilities) is common and typically in the range of 0.2 to 0.4 for a wide variety of traits. As nicely reviewed by Johnson and Barton (2005), neither of two general classes of theoretical models (mutation-selection and balancing selection) proposed to account for standing variation adequately account for the observed levels of variation with realistic estimates of mutation rates. As Johnson and Barton point out, the resolution of these issues likely resides in more detailed considerations of pleiotropy, wherein new mutations influence a number of traits. Once again, it is likely we need to move to a higher dimensional space to reasonably account for observations based on a projection into one dimension (i.e., standing heritability levels for a trait).

The final consideration when with pleiotropy is not just the higher-dimensional fitness surface for the vector of traits they influence but also the distributional space of pleiotropic mutations themselves. Is the covariance structure $\mathbf{G}$ itself some optimal configuration for certain sets of highly-correlated traits? Indeed, has there been selection on developmental processes to facilitate morphological integration (the various units of a complex trait functioning smoothly together), which in turn would result in constraints on the pattern of accessible mutations under pleiotropy (Olson and Miller 1958, Lande 1980)? Our spherical cow may in reality have a very non-spherical distribution of new mutation phenotypes around a current phenotype.

Closing Comments: G is Dead, Long Live G

In a recent evaluation of the evolutionary quantitative genetics research program, Pigliucci (2006) claimed that the field needed "a fundamental reconsideration of its goals", particularly with regard to the utility of $\mathbf{G}$ as an informative construct. Pigliucci reminds us of three classic limitations of quantitative genetics; the environment-dependent nature of its parameters, the summative nature of those same parameters, and the temptation to use such parameters in making evolutionary inferences based solely on correlational evidence. Rather than following Pigliucci in advocating the abandonment of unresolved questions, we would argue that a major focus of future efforts is escaping the Flatlandian prison in which much of biology finds itself captive. It is perhaps ironic given Pigliucci's dismal of the usefulness of $\mathbf{G}$, that it is an understanding of the geometry of $\mathbf{G}$

from both the point of view of mutation generating genetic variance and from selection depleting it, which may result in a more complete understanding of those basic observations concerning genetic variation and selection that our field has strived to explain for so long.

It would be easy for empiricists to take a depressing message away from our discussion; without being able to measure all the traits that are under selection, the tyranny of Flatland and it's potential for generating incorrect interpretations is ever present. To some extent this is likely always to be true, but are there any mitigating circumstances that make the task more manageable than this message might suggest? It's difficult to say, and a lack of theory that models sufficient dimensions is a hindrance here. Although the investigation of a modest number of dimensions in recent experiments has made some progress in finding associations between selection and the distribution of genetic variance, Pigliucci's (2006) admonishment concerning a reliance on correlational evidence is well made; empirical research programs need to incorporate either manipulations to test the associations found, or sampling designs from natural populations which allow clear tests of the evolutionary inferences that are being made.

Finally, there are very deep issues at the core of whether depletion of genetic variation along certain orientations, while still maintaining significant variance in all component directions, is a common phenomenon. The heart of the problem is the nature of the maintenance of genetic variation. We see an explosion of variance in just about every (univariate) trait we examine, values too high for standard models to account for. However, Fisher's original model (spherical cow or not) may provide a key bit of insight. Genetic variation may indeed be ruthlessly exploited by persistent selection along a particular direction, with further progress (adaptation) dependent upon adaptive new mutations. As Fisher showed, improving fitness in a high dimensional space is not trivial, because while a mutant could have enhanced fitness along several dimensions, it may have reduced fitness along others. The net result may be equivalent to the nearly-neutral models from molecular population genetics (Kimura 1983), wherein mutation is pumping into the population new variants that are close to effectively neutral, but not too close. Thus, their variation is higher than expected under standard mutation-selection models (as they have weak selection) but far less than expected under pure drift-mutation models (as they have some selection). Perhaps this is what we are observing with most quantitative variation. In essence, we may have a misleading view; due to projections of variation on single traits, there appears to be considerable genetic variation to exploit. There may well be, *provided* we are trying to move in a different direction relative to persistent selection. If we attempt to improve along lines of historically persistent selection, very little gain may be the expectation.

Acknowledgments

We are extremely pleased to dedicate this paper to Stuart Barker. For someone learning the trade of quantitative genetics in the late 1980's, Stuart's work was like a beacon of interest in a sea of allozymes; incisive reviews, classic experimental designs (even with allozymes!), and above all the innovative application of quantitative genetics to important and interesting questions in evolutionary biology. MWB would like to acknowledge his close collaborators on many aspects of this work, Steve Chenoweth, Emma Hine and Katrina McGuigan. Finally, we thank both reviewers for very helpful comments.

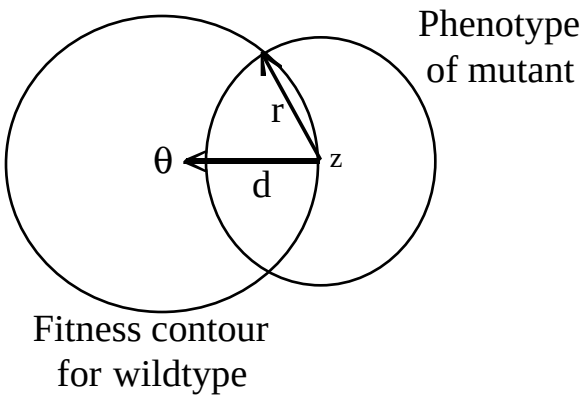

Figure 1: Fisher's model of the chance that a new a mutation is adaptive (shown here for the two dimensional case). The current phenotype is at point z, while the optimum of this fitness surface is at point θ, which is (Euclidean) distance d away from the current phenotype. A random mutation creates a new phenotype at distance r from the wildtype. The circle centered at θ, and passing thorough z, represents the equal-fitness contour of the current phenotype. The circle with diameter r centered at z represents the potential range of new mutation given a move of r in phenotypic space. The probability of adaptation is simply that part of this circle that is inside (i.e., closer to θ) of the current fitness contour of the wildtype. In this case, roughly (by eyeball) 30%.

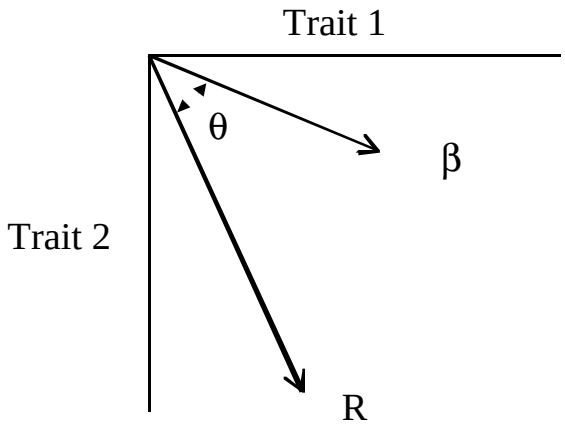

Figure 2: Mild constraints can be imposed even for a diagonal $\mathbf{G}$ matrix (i.e., no genetic covariances between traits). Unless the pattern of genetic variances corresponds to the strength of selection ($\beta_i = c \cdot G_{ii}$), then $\boldsymbol{\beta}$ is not an eigenvector of $\mathbf{G}$, and the actual response $\mathbf{R}$ will be rotated somewhat away from the optimal response $\boldsymbol{\beta}$. These two vectors are plotted for our example where $\beta_1 = 2, \beta_2 = -1, G_{11} = 10, G_{22} = 40, G_{12} = 0$. The resulting angle θ between $\mathbf{R}$ and $\boldsymbol{\beta}$ is 37 degrees.

References

Abbott, E. A. 1884. *Flatland. A romance of many dimesnions.* Seeley and Co. Ltd., London. (Reprinted in 1922 by Dover Publications).

Amemiya, Y. 1985. What should be done when an estimated between-group covariance-matrix is not nonnegative definite. *Am. Stat.* 39: 112–117.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macro-evolution. *Genetica* 112: 9–32.

Barton, N. H. 1990. Pleiotropic models of quantitative variation. *Genetics* 124: 777–782.

Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and fitness surface for multiple male sexually selected traits. *Amer. Natl.* 163: 329–340.

Bohren, B. B., W. G. Hill, and A. Robertson. 1966. Some observations on asymmetical correlated responses to selection. *Genet. Res. Camb.* 7: 44–57.

Bulmer, M. G. 1971. The effect of selection on genetic variability. *Amer. Nat.* 105: 201–211

Dickerson, G. E. 1955. Genetic slippage in response to selection for multiple objectives. *Cold Spring Harbor Symp. Quant. BIol.* 20: 213–224.

Fisher, R. A. 1930. *The genetical theory of natural selection.* Oxford University Press, Oxford, U. K.

Hartl, D., and C. H. Taubes. 1996. Compensatory nearly neutral mutations: selection without adaptation. *J. Theor. Biol.* 182: 303–309.

Hazel, L. N. 1943. The genetic basis for constructing selection indexes. *Genetics* 28: 476–490.

Henderson, C. R., and R. L. Quaas. 1976. Multiple trait evaluation using relatives' records. *J. Anim. Sci.* 43: 1188–1197.

Hill, W. G., and R. Thompson. 1978. Probabilities of non-positive definite between-group or genetic covariance matrices. *Biometrics* 34: 429–439.

Hine, E., and M. W. Blows. 2006. Determining the effective dimesnionality of the genetic variance-covariance matrix. *Genetics* 173: 1135–1144.

Johnson, T., and N. Barton. 2005. Theoretical models of selection and mutation on quantitative traits. *Phil. Trans. R. Soc. B* 360: 1411–1425.

Kimura, M. 1983. *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge, U. K.

Kirkpatrick, M., and K. Meyer. 2004. Direct estimation of genetic principal components: simplied analysis of complex phenotypes. *Genetics* 168: 2295–2306.

Kruuk, L. E. B., J. Slate, J. M. Pemberton, S. Brotherstone, F. Guinness, and T. Clutton-Brock. 2002. Antler size in red deer: heritability and selection but no evolution. *Evolution* 56: 1683–1695.

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

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

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

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

Leigh, E. G. 1987. Ronald Fisher and the development of evolutionary theory. II. Influences of new varation on the evolutionary process. *In*, P. H. Harvey and L. Partridge (Eds) *Oxford Surveys in Evolutionary Biology* 4: 213–263.

Lerner, I. M. 1950. *Population Genetics and Animal Improvement.* Cambridge University Press.

Lush, J. L. 1948. *The Genetcios f Populations.* Iowa State University, Ames.

McGuigan, K., and M. W. Blows. 2007. The phenotypic and genetic covariance structure of Drosphilid wings. *Evolution* 61: 902-–911.

McGuigan, K., S. F. Chenoweth, and M. W. Blows. 2005. Phenotypic divergence along lines of genetic variance. *Amer. Nat.* 165: 32-43.

Merilä, J., B. C. Sheldon, L. E. B. Kruuk. 2001. Explaining stasis: microevolutionary studies in natural populations. *Genetica* 112-113: 199–222.

Meyer, K., and M. Kirkpatrick. 2005. Restricted maximum likelihood estimation of genetic principal components and smoothed covariance matrices. *Genet. Sel. Evol* 37: 1-30.

Mezey, J. G., and D. Houle. 2005. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution* 59: 1027–1038.

Olson, E., and R. Miller. 1958. *Morphological Integration.* University of Chicago Press, Chicago.

Orr, H. A. 1988. The population genetics of adaptation: the distirbution of factors fixed during adaptive evolution. *Evolution* 52: 935–949.

Orr, H. A. 2000. Adaptation and the cost of complexity. *Evolution* 54: 13–20.

Pearson, K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Phil. Trans. Roy. Soc. Lond.* A 200: 1–66.

Pease, C. M., and J. J. Bull. 1988. A critique of methods for measuring life history trade-offs. *J. Evol. Biol.* 1: 293–303.

Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. *Evolution* 43: 1209–1222.

Pigliucci, M. 2006. Genetic variance-covariance matrices: a critique of the evolutionary quantitative genetics research program. *Biology and Philosophy* 21: 1–23.

Price, G. R. 1970. Selection and covariance. *Nature* 227: 520-521.

Prout, T. and J. S. F. Barker. 1993. F Statistics in *Drosophila buzzatii*: Selection, Population Size and Inbreeding *Genetics* 134: 369–375.

Rice, S. H. 1990. A geometric model for the evolution of development. *J. Theor. Biol.* 143: 319–342.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. *Anim. Prod.* 8: 95–108.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50: 1766–1774.

Shaw, F. H., R. G. Shaw, G. S. Wilkinson, and M. Turelli. 1995. Changes in genetic variances and covariances: G whiz! *Evolution* 49: 1260–1267.

Sheridan, A. K., J. S. F. Barker. 1974. Two-trait selection and the genetic correlation. II. Changes in the genetic correlation during two-trait selection. *AUst. J. Biol. Sci.* 27: 89–101.

Simpson, G.G. 1944. *Temp and mode in evolution.* Columbia University Press, New York, New York.

Smith, H. F. 1936. A discriminant function for plant selection. *Ann. Eugen.* 7: 240–250.

Spitze, K. 1993. Population Structure in *Daphnia obtusa*: Quantitative Genetic and Allozymic Variation *Genetics* 135: 367–374.

Tallis, G. M. and P. Leppard. 1988. The joint effects of selection and assortative mating on multiple polygenic characters. *Theor. Appl. Genet.* 75: 278–281.

Thompson, D. W. 1917. *On Growth and Form.*

Van Homrigh, A., M. Higgie, K. McGuigan, and M. W. Blows. 2007. The depletion of genetic variance by sexual selection. *Current Biology* 17: 528–532.

van Valen, L. 1973. A new evolutionary law. *Evol. Theory* 1: 1–30.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proc. VI International Congress of Genetics* 1: 356–366.

Zeng, Z. B. 1988. Long-Term Correlated Response, Interpopulation Covariation, and Interspecific Allometry. *Evolution* 42: 363–374.