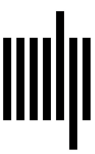
Evolvability

A Unifying Concept in Evolutionary Biology?

Edited by: Thomas F. Hansen, David Houle, Mihaela Pavličev, Christophe Pélabon

Citation:

Evolvability: A Unifying Concept in Evolutionary Biology?
Edited by: Thomas F. Hansen, David Houle, Mihaela Pavličev, Christophe Pélabon
DOI: 10.7551/mitpress/14126.001.0001
ISBN (electronic): 9780262374699
Publisher: The MIT Press
Published: 2023



The MIT Press

5 Variation, Inheritance, and Evolution: A Primer on Evolutionary Quantitative Genetics

Thomas F. Hansen

Evolutionary quantitative genetics (EQG) emerged as a research paradigm in the 1980s based on operational tools for studying variation, inheritance, and selection in field and lab studies. In this chapter, I review the conceptual foundations of EQG as well as newer developments, with particular emphasis on the representation of evolvability and constraints.

5.1 Introduction

Quantitative genetics grew out of the biometry initiated by Francis Galton in the late 1800s. Galton wanted a quantitative science of inheritance and variation, and he was the first to systematically collect and analyze data on variation and inheritance in human populations (Bulmer 2003). He provided the first quantitative evidence for resemblance of offspring to their parents, and thereby for a central part of Darwin's theory of evolution by natural selection. Unfortunately, Galton made mistakes in the interpretation of his findings, which may have contributed to the long delay in the acceptance of natural selection on quantitative variation as the main force of evolution. Galton observed that the offspring of extreme (e.g., selected) parents tended to be more similar than their parents to the population average and took this to be a force working against selection. In Galton's model, selection could only make transient changes that would regress back toward the starting point. This led him and others to the view that selection on "biometric" variation was impotent and irrelevant to macroevolution, which instead must happen through discrete qualitative "mutations" that allow permanent change in the species (Provine 1971).

These misunderstandings were corrected when it was realized that biometric variation was based on discrete Mendelian genes. Even if offspring deviate less than their parents from the population mean on average (the heritability is less than unity), they are still different from the mean, and in the Mendelian model of inheritance, this difference is permanent and will not regress. This is captured in the breeder's equation: $\Delta\bar{z} = h^2 S$, where the change in the mean of a trait, z, from one generation to the next is equal to the heritability, h^2, multiplied with the selection differential, S, the average difference between selected parents and the population mean. By the time Fisher wrote his foundational paper on quantitative genetics in 1918, it was becoming clear that quantitative variation was compatible with a genetic architecture consisting of many Mendelian factors of small effect, and that selection on such variation could produce nonregressive open-ended evolution.

Importantly, it was understood that recombination between Mendelian factors could produce phenotypes that were far outside the observed range of variation, and that selection could thus do more than picking out the most extreme "lineage", as was commonly thought at the turn of the century (Beatty 2016). Fisher's model also revealed that only a part of the genetic variation, the so-called additive component, is reliably inherited during a response to selection

These insights are the foundation of quantitative genetics, but they did not earn the field a central place in the modern synthesis. Instead, the reduction of biometric variation to Mendelian genetics seems to have been taken as a license to focus on one- or two-locus allele-frequency dynamics, while leaving implicit the evolution of complex phenotypes. Quantitative genetics survived and developed in the applied breeding sciences, but for the next 50 years, it was less important to the development and application of evolutionary theory.

This situation changed in the 1980s with the emergence of a new evolutionary quantitative genetics (EQG), which soon became a major driver of empirical research in evolutionary biology. The emergence of EQG happened in two steps. The first was a set of theoretical papers by Russ Lande in the late 1970s (e.g., Lande 1976a,b, 1979, 1980). In these papers, Lande derived equations describing the evolution and maintenance of variation in quantitative traits under selection, drift, and mutation. He hypothesized that genetic variation in polygenic traits was stable and able to support open-ended evolution on macroevolutionary time scales. The second step was the development of operational measures of selection and constraints that provided a framework for empirical studies of the microevolutionary process. In particular, Lande and Arnold (1983) presented the selection-gradient approach for studying selection with multiple regression techniques. This allowed simple and effective ways of separating direct and indirect selection, and studying modes and causes of selection (e.g., Arnold 1983; Arnold and Wade 1984a,b; Wade and Kalisz 1990).

The focal point of EQG is the *Lande equation,* $\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$, which expresses the response to selection of a multivariate (vector) trait, $\mathbf{z}$, as a product of the G-matrix, $\mathbf{G}$, describing additive genetic variation in the trait vector and the selection gradient, $\boldsymbol{\beta}$, describing directional selection on the traits. This provides a neat separation between evolvability on one side and selection on the other. Each of these can be studied statistically in isolation and then be brought together for interpretation in a common theoretical framework. This heuristic generated two interlinked empirical research programs: One centered on field studies of selection based on the selection gradient, and the other centered on lab and field studies of genetic constraints and evolvability based on the G-matrix.

In this chapter, I first explain the statistical model of the genotype-phenotype map initiated by Fisher in 1918 before I present the basic theory of EQG with emphasis on the operationalization of selection and evolvability. I focus on basic principles, concepts, and measurement, and I do not review the many empirical findings of EQG research over the past 40 years. Although a comprehensive overview of these findings is not available, some elements can be found in Charmantier et al. (2014) and Hansen and Pélabon (2021).

5.2 The Statistical-Genetics Model of Variation and Inheritance

At the core of quantitative genetics sits a statistical model of the genotype-phenotype map that was introduced by Fisher (1918). To understand this model as it stands today (e.g., Lynch and Walsh 1998), imagine that we were asked to predict the stature of a random

individual, which we will take to be Francis Galton. If we have no knowledge about Galton, except perhaps that he is a male, the best we can do is to guess that he is at the (male) average of the population. If we denote stature with z, and the population average is $\bar{z} = 170\ cm$, our initial guess for Galton's stature is

$$z = \bar{z} = 170\ cm.$$

Now, say we learn that Galton was a carrier of a specific allele at the locus *ZBTB38*, which codes for a transcription factor that is involved in the regulation of the thyroxin hydroxylase gene. This locus was found to have one of the largest effects on human stature in the Genome-Wide-Association study (GWAS) of Gudbjartsson et al. (2008). Carriers of the specific allele are on average about $0.5\ cm$ taller than the population at large. With this information, we predict that the stature of Galton is

$$z = \bar{z} + \alpha_{zbtb38} = 170\ cm + 0.5\ cm = 170.5\ cm.$$

The entity α is the average deviation (excess) of the allele. Such deviations can be defined for all the alleles that segregate in the population, and if we knew the average deviations of all the alleles in Galton's genome, we could sum them together to get a better prediction of his stature. The real Galton was tall, so say that the sum is a positive $10\ cm$; then our prediction is

$$z = \bar{z} + \sum_i \alpha_i = 170\ cm + 10\ cm = 180\ cm,$$

where the sum is over all alleles in the given genome (i.e., twice the number of loci for a diploid organism). This sum, $A = \sum_i \alpha_i$ is known as the breeding value of the individual, because with some assumptions, we can predict the average trait of offspring from the average of the breeding values of their parents. The breeding value of Galton is thus $A = 10\ cm$. For simplicity, I will also refer to A and the α as additive effects.

Knowing the additive effect of his genes gives us a better prediction of Galton's height but is still inaccurate, because we have no biological basis for our assumption that the average deviations can be added together. For example, imagine that Galton had not one, but two copies of the tall allele on the *ZBTB38* locus. Then this should add $1\ cm$ to his predicted height, but if it was sufficient with only one copy for the biological effect, this prediction would be wrong. We can account for this possibility by adding a dominance deviation, defined as the average deviance of the phenotype of the carriers of the two alleles from the additive prediction based on the two alleles. If it did not matter whether there were one or two tall *ZBTB38* alleles, their dominance deviation would be $\delta = -0.5\ cm$. We can calculate such dominance deviations for the allele pairs at all Galton's loci, and if we add them together we get a dominance effect $D = \sum_i \delta_i$. If we pretend that $D = -5\ cm$ for Galton, we get

$$z = \bar{z} + A + D = 170\ cm + 10\ cm - 5\ cm = 175\ cm,$$

as a better prediction of his stature. Likewise, there is no biological reason for adding effects across loci, and for each pair of alleles at different loci, call them i and j, we can define an additive-by-additive epistatic deviation $\alpha\alpha_{ij}$ as the average deviation of their carriers from the additive prediction. The sum of these effects over all pairs of alleles is $AA = \sum_i \sum_j \alpha\alpha_{ij}$. Adding this term will account for pairwise deviations from the additive prediction but not higher-order interactions. It may be that individuals who carry three specific alleles will

A_1	B_1
A_2	B_2

$$
\begin{aligned}
z = \bar{z} &+ \alpha_{A1} + \alpha_{A2} + \alpha_{B1} + \alpha_{B2} && (A)\\
&+ \delta_{A1A2} + \delta_{B1B2} && (D)\\
&+ \alpha\alpha_{A1B1} + \alpha\alpha_{A1B2} + \alpha\alpha_{A2B1} + \alpha\alpha_{A2B2} && (AA)\\
&+ \alpha\delta_{A1A2B1} + \alpha\delta_{A1A2B2} + \alpha\delta_{A1B1B2} + \alpha\delta_{A2B1B2} && (AD)\\
&+ \delta\delta_{A1A2B1B2} && (DD)\\
z = \bar{z} &+ A + D + AA + AD + DD
\end{aligned}
$$

Figure 5.1
The statistical-genetics model for a trait determined by two loci: The genetic value of an individual with alleles A_1 and A_2 at one locus and B_1 and B_2 at the other locus is decomposed into all possible component effects. The four first-order additive effects are given on the first line, the two dominance effects on the second line, the four pairwise *AA* epistatic effects on the third line, the four third-order *AD* epistatic effects on the fourth line, and the single fourth-order *DD* epistatic effect is given on the fifth line.

deviate on average from the prediction based on the additive, dominance, and pairwise epistatic effects involving those three alleles. This deviation can then be added to make an even better prediction. Such average deviations from lower-order predictions can be calculated for any set of alleles, and if we could calculate and add together the effects of all possible subsets of alleles from Galton's genome, we would get a perfect prediction of his total genotypic value, *G*. In figure 5.1 this is illustrated for a trait influenced by two loci. In addition to the *AA* effect, we recognize *AD* effects due to interactions among a pair of alleles at one locus and a single allele at another locus and *DD* effects due to interactions among all four alleles at two loci. With more loci, there are also *AAA* effects due to deviations of triplets of alleles at different loci and higher-order effects.

Adding all these effects together yields the total epistatic effect, *I*. If Galton's total epistatic value is $I = AA + AD + DD + AAA + (\text{higher-order effects}) = 3\ cm$, we get

$$z = \bar{z} + A + D + I = 170\ cm + 10\ cm - 5\ cm + 3\ cm = 178\ cm,$$

and Galton's total genetic value is $G = A + D + I = 8\ cm$. We predict that his genes would make him 8 *cm* taller than the population average, but this is still not a perfect prediction, because it leaves out effects of the environment. Just as each individual has a unique set of alleles, each individual experiences a unique set of environmental influences in terms of different diets, diseases, amounts of stress, and so forth. For each specific environmental influence, one can define an average deviation from the population mean of individuals experiencing this factor, and these deviations can then be summed up to an environmental value, *E*, which can be added to the genetic value. Belonging to the gentry, Galton likely experienced an environment favorable to growth, so we assign him a positive score, say, $E = 6\ cm$. With this we arrive at the prediction

$$z = \bar{z} + G + E = 170\ cm + 8\ cm + 6\ cm = 184\ cm,$$

which should be close to Galton's true height, but just as genes may interact in their effects, some combinations of genes and environments may have effects that deviate from the sum of their separate effects. For example, an individual experiencing a calcium deficit in his diet may be disproportionally affected if he also carries alleles that weakens his calcium uptake. Such deviations are called genotype-by-environment, $G \times E$, interactions, and they

can be summed and added to the genotypic and environmental values. If Galton's $G \times E$ value is $-1\ cm$, we get

$$z = \bar{z} + G + E + G \times E = 170\ cm + 8\ cm + 6\ cm - 1\ cm = 183\ cm,$$

which would be a perfect prediction of Galton's stature, provided we have included and perfectly measured all genetic and environmental effects.

It is not even remotely possible to measure all relevant effects, but this is not the point. The model is useful because it can be combined with (Mendelian) rules of inheritance and segregation of alleles to understand trait variation and inheritance. The population variance in the phenotype, V_P, can be decomposed into contributions from each of the components discussed above by assuming that they are statistically independent of one another:

$$V_P = V_G + V_E + V_{G\times E} = V_A + V_D + V_I + V_E + V_{G\times E},$$

and each of these can be decomposed further. For example, $V_I = V_{AA} + V_{AD} + V_{DD} + V_{AAA} +$ higher-order effects. Also, by assuming Hardy-Weinberg and linkage equilibrium, which is equivalent to assuming that all alleles segregate independently, we can decompose each variance component into contributions from each average deviation. For example, the additive genetic variance is the sum of the variances of all the average deviations of alleles, $V_A = \sum_i \mathrm{Var}[\alpha_i]$, where $\mathrm{Var}[\alpha_i]$ is the population variance in the average deviation of single alleles. The assumption that all alleles segregate independently may seem severe, but covariances among alleles and genotypes can be added to the decomposition or incorporated into the definition of the variance components. The additive variance is usually understood as including these covariances, and the sum of the variances of the average deviations is technically known as the "genic" variance.

Even though all the statistical variance components contribute to differences among individuals, they are not inherited in the same way. Figure 5.2 illustrates the inheritance of the different components with a two-locus model. Our individual from figure 5.1 is mated to another individual, and an offspring is produced by picking one allele from each locus from each parent for the offspring's genotype. Comparing the genotypes of the mother and her offspring reveals that only a minority of the statistical components determining the mother's genetic value are to be found in her offspring. The two share half of their additive deviations, but only one out of the many dominance and epistatic deviations. From the perspective of the offspring, all its additive deviations are inherited from its two parents, but only half of its *AA* epistatic deviations and none of the deviations involving dominance in any form. The reason for the latter is that an individual can never inherit two alleles at one locus from the same parent. Its dominance deviations are therefore different from those of its parents. An exception occurs if the two parents are related. If they are, they may carry some alleles that are identical by descent, and then the offspring may end up with the same two allele copies as a parent at some loci (see Lynch and Walsh 1998 for details).

On the population level, these observations imply that all the additive genetic deviations in a population have been inherited (in a rearranged form) from the parental population, but none of the dominance deviations and only fractions of epistatic deviations are inherited. The noninherited genetic deviations have been generated de novo by recombination. This has consequences for evolution by natural selection. Selection acts similarly on all

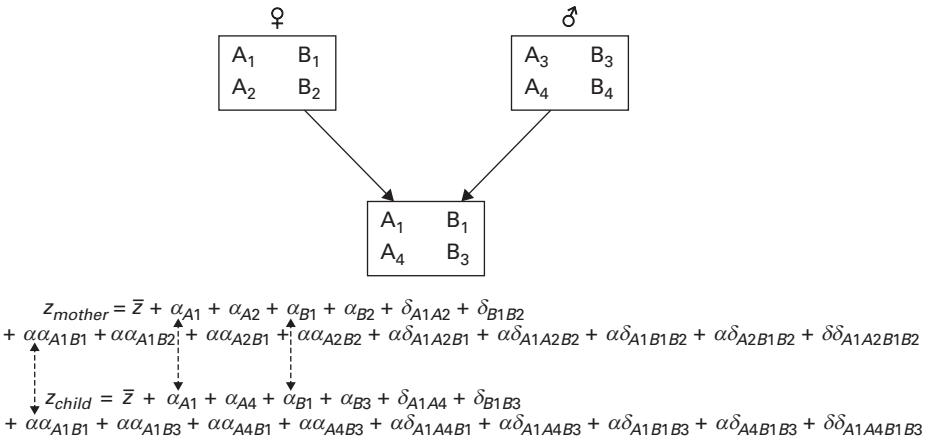


Figure 5.2
Inheritance of a two-locus trait: Two individuals mate to produce an offspring, which inherits the alleles A_1 and B_1 from the mother, and the alleles A_4 and B_3 from the father. The full genotypic values of the mother and the offspring are spelled out. The dashed arrows indicate the shared effects. These are two additive effects, α_{A1} and α_{B1}, and one epistatic effect, $\alpha\alpha_{A1B1}$. This shows that half the additive effects and one quarter of the *AA* epistatic effects are shared between a parent and the offspring. Also note that all the additive effects of the offspring are inherited from the two parents, while none of the dominance effects and only fractions of the epistatic effects are inherited.

components, but only selection on additive variation is fully transmitted to the next generation. Although some epistatic deviations are also inherited, and indeed contribute to the selection response, this response is transient in the sense that it is broken down by recombination.

For these reasons, it is the additive component of variance that determines the short-term evolvability of a population, which explains the focus on the additive variance in EQG. In particular, it explains the central role of the additive genetic variance matrix, **G**, which summarizes the additive genetic variances and covariances of a set of traits. For a vector, $\mathbf{z} = \{z_1, z_2, z_3\}^T$, of three traits, the G-matrix is

$$\mathbf{G} = \begin{bmatrix} V_{A1} & C_{A12} & C_{A13} \\ C_{A12} & V_{A2} & C_{A23} \\ C_{A13} & C_{A23} & V_{A3} \end{bmatrix},$$

where V_{Ai} is the additive genetic variance of trait i, and C_{Aij} is the additive genetic covariance of trait i and j. Variance matrices describe variances of vectors, and the phenotypic variance matrix, denoted **P**, can be decomposed into additive genetic, dominance, epistatic, and environmental components just as described for the variance of a univariate trait. The inheritance also works in the same way. A covariance can be thought of as the amount of variance that is shared between two variables, and just as for variances, it is the additive part that is stably inherited.

The statistical model also provides the tools for estimation of the different variance components. By comparing the genetic makeup of relatives, as was done for mother and offspring in figure 5.2, it is easy to compute what they share. In the case of parent and offspring, we saw that they share half the additive variance, a quarter of the additive-by-additive variance, and none of the dominance variance. Similar computations can be done

for other types of relatives. Fullsibs also share half of the additive variance and a quarter of the additive-by-additive variance, but additionally, they share a quarter of the dominance variance. Excess similarity (i.e., covariance) of fullsibs compared to parent with offspring can thus be used to estimate dominance variance. Halfsibs are particularly well suited to estimate additive variance, because they share a quarter of the additive variance but no dominance and only one sixteenth of the additive-by-additive variance. More generally, known or estimated pedigrees can be used to find the shared components of any pair of individuals in a population, and this can be used in a mixed-model framework to estimate the different components of variance as well as other aspects of genetic architecture (Lynch and Walsh 1998; Wilson et al. 2010; Jensen et al. 2014; Morrissey et al. 2014).

5.3 Selection in EQG

Natural selection occurs when individuals in a population have different properties, and these properties affect the ability of the individuals to survive and reproduce. If the properties are heritable in the sense that offspring are (statistically) similar to their parents, then we have evolution by natural selection. Evolution by natural selection can thus be seen as a two-step process with a first selection step and a second transmission step, in which the properties of the selected parents are transmitted to their offspring. Although transmission can be complicated, the mathematics of selection is simple.

An episode of selection can be described mathematically as a mapping from a set of individuals before selection to another set of individuals after selection (Price 1970; Arnold and Wade 1984a; Kerr and Godfrey-Smith 2009). Since the goal is to understand the effects of selection on some property, the individuals are classified into types that share the property of interest (e.g., Otsuka 2019). The types may be speckled and melanic moths, as in the classical textbook example of selection due to industrial pollution (Majerus 1998); all individuals that share a particular allele or genotype, as in classical population genetics; or all individuals that share a particular value of a quantitative trait, say a stature of 183 *cm*. The key point is that each type is assigned a fitness, a number that represents its contribution to the set after selection. In biological terms, the fitness is the ability of the type to survive and reproduce, and in mathematical terms, it is a random variable with a type-dependent distribution across individuals (Hansen 2017). In practice, we consider only the expected value of the fitness for a type (as in the propensity interpretation of fitness). Formally, the fitness of a type, z, is defined as the ratio of the amount of the type after selection, N_z', and the amount before selection, N_z. This yields the relation

$$N_z' = W(z)N_z,$$

where $W(z)$ is the absolute fitness of type z. Usually, the interest is not in the absolute amount of the type, but in its frequency relative to other types. If p_z is the frequency (fraction of total amount) of type z before selection, then some algebra yields its frequency after selection:

$$p_z' = w(z)p_z,$$

where $w(z) = W(z)/\overline{W}$, with $\overline{W}$ the mean fitness of the population. The $w(z)$ is the relative fitness of the type. Hence, absolute fitnesses describe changes in amount of types, and relative fitnesses describe changes in the frequency of types. From this equation, we can

also calculate that the change in the mean trait value due to selection alone (i.e., the selection differential S) equals the covariance between the trait value and relative fitness:

$$S = Cov[w(z), z].$$

Price (1970) added transmission to this equation:

$$\Delta\bar{z} = Cov[w(z), z] + \overline{w(z)(z'(z) - z)},$$

where $z'(z)$ is the expected trait value of offspring of parents with trait value z, and the average in the transmission term is taken over the population distribution of z. The whole transmission term can be interpreted as the average difference in trait value between offspring and their parents. It is weighted with fitness to account for some types contributing more offspring than others.

The breeder's equation can be derived from the Price equation by assuming that the parent-offspring regression is linear with slope equal to the heritability, h^2. Then it follows from simple geometry that the average difference between offspring and their parents is $(h^2 - 1)S$, and we get

$$\Delta\bar{z} = S + (h^2 - 1)S = h^2 S.$$

Although the breeder's equation has been the standard representation of the response to selection in quantitative genetics, the foundation of EQG was based on a conceptually important rearrangement of the equation. I refer to this rearrangement as the Lande equation, and in the univariate case, it goes as follows:

$$\Delta\bar{z} = h^2 S = \frac{V_A}{V_p} S = V_A \frac{S}{V_P} = V_A \beta,$$

where $\beta = S/V_P$ is the selection gradient (i.e., the linear regression slope of relative fitness on the trait). Given this and defining heritability as $h^2 = V_A/V_P$, the Lande and the breeder's equations are mathematically equivalent. They are not conceptually equivalent, however, because they make different assumptions about what entities go together as quasi-independent units. In the case of the breeder's equation, these entities are the heritability and the selection differential; and in the case of the Lande equation, they are the additive variance and the selection gradient. Choices of which conceptual entities to measure and use in models constrain our thinking and interpretation of results, and the Lande equation paved the way for the study of evolvability separate from selection in a way that is not apparent with the breeder's equation (see section 5.8).

An instructive alternative derivation of the Lande equation is to assume that the breeding value, A, is the only component of the genotype that is transferred from parents to offspring, and that the breeding value of the offspring is equal in expectation to the breeding value of the parents. Then the evolutionary change in the trait equals the evolutionary change in the breeding value, and we get

$$\Delta\bar{z} = \Delta\bar{A} = Cov[w(z), A] = V_A \frac{Cov[w(z), A]}{V_A} = V_A \beta_A,$$

where β_A is the selection gradient on the breeding value of the trait. Hence, we see that the Lande equation also follows from the assumptions that only the breeding value is

transmitted and that the phenotypic selection gradient, β, equals the (additive) genetic selection gradient, β_A. The latter is not a trivial assumption, because selection on phenotype and genotype may differ (e.g., Morrissey et al. 2010; Reid and Sardell 2011).

Lande (1979) defined the selection gradient theoretically in terms of the derivative (gradient) of relative fitness with respect to the trait (vector), but Lande and Arnold (1983) proposed to study it as a statistical regression of relative fitness on the trait (vector). The regression approach provides a simple method for studying the trait-fitness relationship in natural populations and is one reason for the success of EQG as an empirical research program. All one needs to use this are measures of traits and realized fitnesses of individuals, and then the coefficients from a multiple regression of relative fitness on the traits estimate the elements of the vector selection gradient

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix},$$

shown here for three traits. Here, β_i, the partial selection gradient on trait i, yields the linear selection on trait i controlling for indirect selection due to correlations with other (included) traits. This technique can be used to distinguish direct and indirect selection, and it lends itself to experimental and observational studies in which hypothesized causes of selection can be tested by comparing selection gradients in their presence and absence. In this respect, the Lande-Arnold framework and the selection gradient are superior to the breeder's equation and the selection differential, which does not easily separate direct and indirect selection.

A key step in applying the framework is to find a useful measure of fitness. Most studies use proxy variables called "fitness components" that can be assumed to represent fitness over an episode of selection when other factors are kept constant. The trick is to find a fitness component that captures the causal mechanism under investigation. In a study of sexual selection on a mating display, one may, for example, measure the number of matings an individual obtains during a mating season. This will not capture all selection on the trait (e.g., through survival), but it may reasonably capture the influence of sexual selection.

Selection gradients are tools to describe fitness landscapes. The selection gradient points in the direction of steepest ascent in the fitness landscape, and its length is a measure of how steep the landscape is in this direction. The response to selection may not follow the selection gradient, however. As illustrated in figure 5.3, the response is generally deflected toward directions of higher evolvability (Lande 1979; Arnold 1994; Schluter 1996; Hansen and Houle 2008). Information about the curvature of the fitness landscape can also be obtained through second-order selection gradients, which can be estimated by including quadratic and interaction terms in the regression. These are collected in the so-called gamma matrix (e.g., Blows 2007), here illustrated for three traits:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{12} & \gamma_{22} & \gamma_{23} \\ \gamma_{13} & \gamma_{23} & \gamma_{33} \end{bmatrix},$$

where the diagonal γ_{ii} terms are the second derivatives with respect to traits. Negative values indicate stabilizing selection on the trait, and positive values indicate disruptive

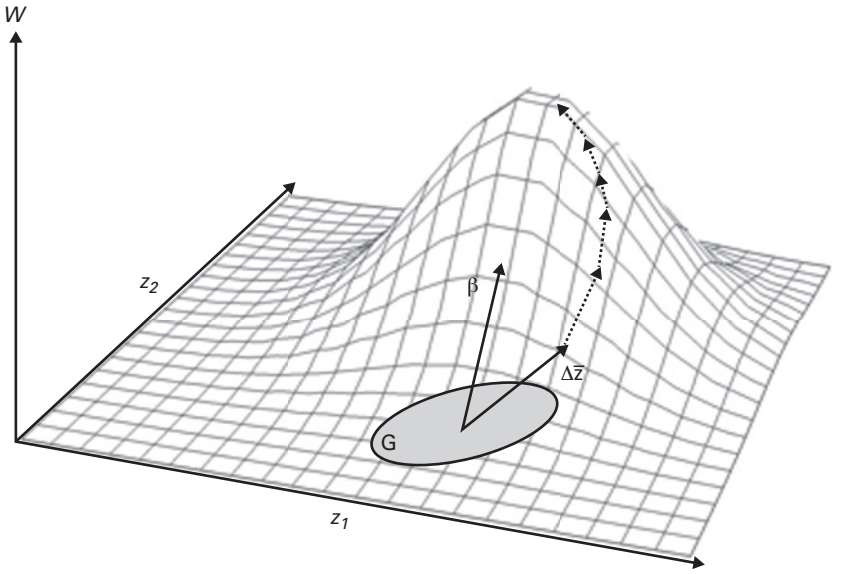


Figure 5.3
Evolution on a fitness landscape: The surface plots fitness against two traits. The ellipse **G** represents the shape of the G-matrix. The selection gradient $\boldsymbol{\beta}$ points from the population mean in the direction of the steepest ascent in the fitness landscape. The response to selection, $\Delta\bar{z}$, is deflected from the selection gradient toward directions of higher evolvability. With a single-peaked landscape, subsequent responses (dotted lines) will curve toward the peak and eventually bring the population mean to peak fitness. In every generation, mean fitness will increase with $e(\boldsymbol{\beta})|\boldsymbol{\beta}|^2$, where $e(\boldsymbol{\beta})$ is evolvability in the direction of the selection gradient and $|\boldsymbol{\beta}|$ is the magnitude of the selection gradient.

selection. Be aware that the quadratic regression coefficients need to be multiplied by 2 to yield the diagonal γ_{ii} (Stinchcombe et al. 2008). The off-diagonal γ_{ij} terms describe correlational selection (i.e., how the selection gradient on one trait changes when another trait is being changed).

5.4 Evolvability and Constraints in EQG

Although the tools for empirical study of selection were one reason for success of EQG, the other reason was the possibility for studying genetic constraints with the G-matrix (e.g., Lande 1979; Cheverud 1984; Arnold 1992). The G-matrix determines how much the selection response is deflected from the selection gradient, and how easy it is to evolve in different directions in morphospace. With three traits, the Lande equation yields the following expression for the response to selection in trait z_1:

$$\Delta\bar{z}_1 = V_{A1}\beta_1 + C_{A12}\beta_2 + C_{A13}\beta_3.$$

The first term is due to direct selection on the trait itself. The other terms show that selection on z_2 and z_3 (i.e., β_2 and β_3) causes a correlated response in our focal trait if there is additive genetic covariance between the traits.

The G-matrix describes genetic constraint (or evolvability) in two ways. The first is in terms of how much additive genetic variance there is in the different traits and trait combinations,

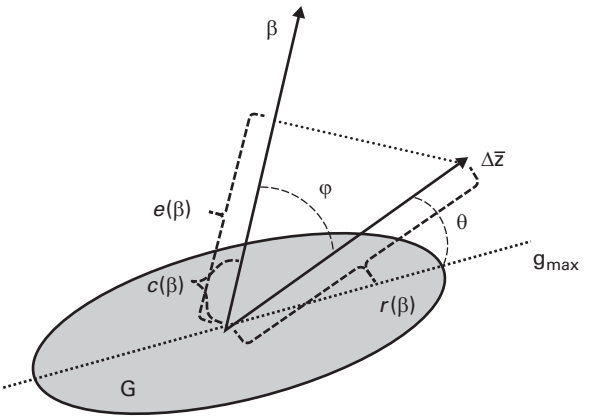


Figure 5.4
The geometry of evolvability and constraint measures. The figure shows the response, $\Delta\bar{\mathbf{z}}$, to a selection gradient, $\boldsymbol{\beta}$, as determined by an additive genetic variance matrix $\mathbf{G}$, with variation distributed as indicated by the ellipse. The major axis of variation in $\mathbf{G}$ is the vector $\mathbf{g}_{max}$, giving the direction of maximum evolvability. Schluter (1996) termed this vector the *line of least genetic resistance* and proposed to use the angle, θ, between this and the response as a measure of (lack of) constraint. The *evolvability,* $e(\boldsymbol{\beta}) = \boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}$, in the direction of the selection gradient is defined as the length of the projection of the response on the gradient. It measures how far the population mean is shifted in the direction of selection. This value is less than the length of the response vector, which is the *respondability,* $r(\boldsymbol{\beta}) = \sqrt{\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}$. The ratio between the evolvability and the respondability equals the cosine of the angle φ between the response and the selection gradient, and it measures how close the response tracks the direction of selection. Marroig et al. (2009) termed this the *flexibility.* The *conditional evolvability* in the direction of the selection gradient, $c(\boldsymbol{\beta}) = (\boldsymbol{\beta}^T\mathbf{G}^{-1}\boldsymbol{\beta})^{-1}$, measures the evolvability if the response is constrained to follow the gradient by selection against deviations, as on an adaptive ridge. The ratio between the conditional and the unconditional evolvability is the *autonomy:* $a(\boldsymbol{\beta}) = c(\boldsymbol{\beta})/e(\boldsymbol{\beta})$. It measures the fraction of genetic variance in direction $\boldsymbol{\beta}$ that is free from constraints. All equations are derived in Hansen and Houle (2008) and require standardizing $\boldsymbol{\beta}$ to unit length. Most of these statistics can be computed with G. Bolstad's R-package "evolvability" (Bolstad et al. 2014).

and hence their capability to respond to selection. The second is that it shows how selection on other traits can interfere with (or enhance) evolution on a focal trait through genetic covariances.

Figure 5.4 illustrates some statistics for describing the potential for evolution conveyed by a G-matrix. First, because evolvability may differ in different directions in phenotype space, Hansen and Houle (2008) proposed a measure of evolvability in a vector direction $\mathbf{z}$ as $e(\mathbf{z}) = \mathbf{z}^T\mathbf{G}\mathbf{z} / |\mathbf{z}|^2$. This statistic has some desirable properties for a multivariate measure of evolvability: (1) It is the additive genetic variance in the direction of the vector, and thus, (2) it reduces to the additive genetic variance for a univariate trait. (3) If measured along a selection gradient, it is the length of the projection of the response (predicted from the Lande equation) on this gradient. It thus measures the ability to respond in the direction of selection and accordingly, (4) it is proportional to the increase in mean fitness due to selection. But note that $e(\mathbf{z})$ is not a measure of the expected magnitude of the selection response. It is a measure of the ability to respond in the direction of selection and would perhaps have been better named adaptability rather than evolvability. On the other hand, it is also (5) proportional to the expected change (variance) in direction $\mathbf{z}$ due to genetic drift.

The evolvability averaged over all directions in morphospace is $\bar{e} = \mathrm{Tr}[\mathbf{G}]/d$, where Tr is the trace operator, and d is the dimensionality (rank) of the G-matrix. One interesting aspect of this average is that it is unaffected by covariances. Intuitively, one might think

that covariances, and particularly negative covariances, would reduce overall evolvability, but this is not the case. Covariances are not general constraints. What they do is to redistribute evolvability by reducing it in some directions and increasing it in other directions.

The $e(\mathbf{z})$ thus describes constraints on the first level as the amount of transmissible variation in specific directions. To measure genetic constraints arising from interfering selection on other traits, Hansen et al. (2003a; Hansen 2003; Hansen and Houle 2008) proposed the conditional evolvability $c(\mathbf{y}|\mathbf{x}) = \mathbf{G}_y - \mathbf{G}_{yx}\mathbf{G}_x^{-1}\mathbf{G}_{xy}$, which gives the evolvability of a trait (vector), $\mathbf{y}$, when another trait (vector), $\mathbf{x}$, is not allowed to change. This is based on a partitioning of the G-matrix of the vector $\mathbf{z} = \{\mathbf{y}^T, \mathbf{x}^T\}^T$ as

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_y & \mathbf{G}_{yx} \\ \mathbf{G}_{xy} & \mathbf{G}_x \end{bmatrix},$$

where $\mathbf{G}_y$ and $\mathbf{G}_x$ are the variance matrices of the vectors $\mathbf{y}$ and $\mathbf{x}$, respectively, and $\mathbf{G}_{yx} = \mathbf{G}_{xy}^T$ is their covariance matrix. The conditional evolvability equals the component of additive variance in a trait vector that is independent of another (defined) trait vector. If there is directional selection on a trait, $\mathbf{y}$, that is correlated with other traits, $\mathbf{x}$, under stabilizing selection, the initial response to selection would be determined by $e(\mathbf{y})$, but indirect selection on the constraining traits would shift them from their optima, inducing counteracting selection that would decrease the response of the focal trait. The response in the focal trait would decay toward a rate given by $c(\mathbf{y}|\mathbf{x})$. If the stabilizing selection is strong, this rate is approached in a handful of generations (Hansen et al. 2019). Conditioning on all orthogonal directions yields the statistics $c(\mathbf{z}) = (\mathbf{z}^T\mathbf{G}^{-1}\mathbf{z}/|\mathbf{z}|^2)^{-1}$, illustrated in figure 5.4, which measures how easy it is to move along an adaptive ridge in direction $\mathbf{z}$. The $c(\mathbf{z})$ is inversely proportional to how much fitness must increase along the ridge to drive a given change.

The conditional evolvability, and derived measures such as autonomy, can be seen as operational measures of modularity, because they quantify the ability of parts to evolve independently of one another. The concept of modularity has long antecedents (e.g., Olson and Miller 1958; Berg 1959), but it became a focus of research in evodevo, where it is treated almost as a synonym of evolvability. The idea is that the evolvability of complex integrated organisms requires organization into functional parts (modules), such as limbs or organs, that can evolve quasi-independently of other parts (Riedl 1977; Cheverud 1984; Wagner and Altenberg 1996; Wagner 2014). Consequently, there appeared many studies of modularity and integration through patterns of trait correlation in G- or P-matrices (reviewed in Melo et al. 2016). This has motivated a variety of methods and statistics for testing and comparing constraints (e.g., Lande 1979; Schluter 1996; Hansen et al. 2003a, 2019; Hansen and Houle 2008; Hohenlohe and Arnold 2008; Marquez 2008; Agrawal and Stinchcombe 2009; Hine et al. 2009; Kirkpatrick 2009; Mitteroecker 2009; Hansen and Voje 2011; Chevin 2013; Bolstad et al. 2014; Grabowski and Porto 2017; Sztepanacz and Houle 2019; Cheng and Houle 2020).

5.5 The Role of Mutation

Studying constraints with the G-matrix is predicated on $\mathbf{G}$ being somewhat constant on relevant time scales. Lande (1976b, 1980) put forward the hypothesis that a constant $\mathbf{G}$ conveying high evolvability could be maintained in a balance between mutation and sta-

bilizing selection. This hypothesis is part of the foundation of EQG and is implicit in much EQG research, but it soon came under debate (e.g., Turelli 1984), and it became clear that its validity depends on poorly known empirical information about patterns of selection, genetic architecture, and mutational input. This debate motivated empirical research, including direct attempts to estimate changes and differences in **G** among different species (Arnold et al. 2008; McGlothlin et al. 2018), and attempts to estimate the rate of mutational input to quantitative traits (Houle and Kondrashov 2006; Halligan and Keightley 2009).

The mutation matrix **M** is defined as the amount of new additive genetic variance that arises in a trait vector in each generation by mutation. Under the Gaussian assumptions favored by Lande, the equilibrium **G** under symmetric stabilizing selection with the mean at the optimum is

$$\hat{\mathbf{G}} = \sqrt{n}\boldsymbol{\Gamma}^{-\frac{1}{2}}\left(\boldsymbol{\Gamma}^{\frac{1}{2}}\mathbf{M}\boldsymbol{\Gamma}^{\frac{1}{2}}\right)^{\frac{1}{2}}\boldsymbol{\Gamma}^{-\frac{1}{2}},$$

where $\boldsymbol{\Gamma}$ is the matrix of second-order selection gradients; n is the (effective) number of loci affecting the trait vector; and the $\boldsymbol{\Gamma}$-matrix is for the genotype, and not the phenotype, and thus potentially dependent on the environmental variance. This equation is derived from equation 28.37c in Walsh and Lynch (2018) based on Lande (1980). In addition to the "Gaussian" assumption of normal distribution of genetic effects at each locus, it assumes weak selection, additivity, equivalent loci, and linkage equilibrium. The equation shows that although **G** tends to increase with input of mutational variance and decrease with the strength of stabilizing selection, the relationship is nonlinear with no simple proportionality with either. This can be seen by calculating the evolvability along eigenvectors, **v**, of $\boldsymbol{\Gamma}$ as $e(\mathbf{v}) = \sqrt{n e_m(\mathbf{v})/\gamma}$, where γ is strength of stabilizing selection (magnitude of corresponding eigenvalue), and $e_m(\mathbf{v}) = \mathbf{v}^T\mathbf{M}\mathbf{v}/|\mathbf{v}|^2$ is the mutational variance in the direction of **v**. Note that the equilibrium depends on genetic architecture beyond the M-matrix in the form of the number of loci affecting the traits. Beyond the Gaussian, weak-selection and additivity assumptions, the equilibrium also depends on details of genetic architecture, such as size, bias, and pleiotropy of mutational effects, as well as on population structure and the exact mode of selection (Walsh and Lynch 2018). Epistasis also affects the maintenance of both additive and nonadditive genetic variance (Hermisson et al. 2003).

Hence, it cannot be assumed that **M** and **G** are equivalent in their effects on evolvability, as is often implicit in verbal discussions. The processes maintaining variation must be considered in addition to the input of variation. Nevertheless, the amount and pattern of mutational input is essential to understanding both short-term evolvability and long-term constraints. Subsequent empirical research has vindicated Lande's hypothesis that mutational input is sufficient to maintain high evolvability in the face of selection. This vindication has come both through direct estimates of **M** and other mutational parameters, and through a growing realization that many quantitative traits are influenced by a very large number of locations in the genome. The latter not only generates a large mutational target size, it also makes the Gaussian assumptions more likely.

Houle (1998) proposed that mutational target size was a major determinant of levels of genetic variation. Traits and trait categories with the potential to be affected by many genomic changes are likely to experience more mutation and to harbor more genetic variation, and this

factor may explain more of the "variation in variation" than differences in strengths of selection. This hypothesis explains why life-history traits and fitness tend to have more additive genetic variation than morphological traits despite presumably stronger selection on the former. Life-history traits, such as survival and fertility, have large mutational target sizes, because they are influenced by most aspects of morphology, physiology, and behavior, and thus "inherit" their mutational input. Consequently, the mutational input to life-history traits can be so large as to be able to replenish standing levels of variation in mere tens of generations (Houle 1998; Halligan and Keightley 2009).

5.6 Genetic Drift in EQG

Adaptive genetic evolution is driven by (direct) natural selection, but many evolutionary changes are not adaptive. We have seen that indirect selection and correlated responses are important, and mutation, recombination, migration, and genetic drift also cause change. Genetic drift refers to changes due to random sampling of alleles in finite populations. For a polygenic trait (vector) with a linear genotype-phenotype map, the change in the mean from generation to generation due to drift is normally distributed with mean zero and a variance matrix equal to $\mathbf{G}/N_e$, where N_e is the effective population size, an inverse measure of the strength of genetic drift (Lande 1976a, 1979). Even if there is no trend to the changes, the trait mean will undergo random fluctuations that shifts it away from the ancestral state. The variance of these fluctuations in a trait, $\mathbf{z}$, will increase linearly with time (in generations) and be proportional to the evolvability $e(\mathbf{z})$ in the direction of the trait vector.

This prediction assumes that the G-matrix stays constant, but another effect of genetic drift is to reduce the genetic variance and thus the G-matrix each generation with an expected factor of $1 - 1/2N_e$. Given an input $\mathbf{M}$ of mutational variance, the change in $\mathbf{G}$ in a "neutral" model (i.e., due to mutation and drift alone) is then $\Delta\mathbf{G} = (1 - 1/2N_e)\mathbf{G} + \mathbf{M}$, and from this, we infer a neutral equilibrium of $\hat{\mathbf{G}} = 2N_e\mathbf{M}$. Hence, if there is no selection, we expect more genetic variation in larger populations, and we expect the G-matrix to be proportional to the M-matrix. Lynch (1990) noted that this result implies that the effective population size will drop out of the variance of trait change from generation to generation, which will be equal to $2\mathbf{M}$. Hence, we expect neutral evolution to be equally fast in small and large populations. This is the quantitative-genetics analog of Kimura's molecular clock, in which the substitution rate depends only on the mutation rate and not on the population size. It implies that the expected variance of change in a trait vector, $\mathbf{z}$, in a neutral model is equal to $2e_m(\mathbf{z})$ per generation with $e_m(\mathbf{z})$, defined in section 5.5, being a mutational evolvability.

Although the drift and neutral models are unrealistic as standalone hypotheses for the evolution of quantitative traits, they yield insights into the process of evolution. Contrary to intuition, the neutral model predicts high rates of evolution on macroevolutionary time scales. Lynch (1990) used estimates of mutational variance to argue that almost all observed changes in the fossil record were too small to be compatible with the neutral model, and of course also with sustained directional selection. Hence, we reach the inescapable conclusion that quantitative traits must be under some form of stabilizing selection on long time scales.

In the presence of selection, genetic drift is not likely to be important on the phenotypic level beyond a few generations (the molecular level is another matter; see Lynch 2007).

One instance of this concerns the role of stochastic peak shifts in evolution. We have seen that selection will bring a population to the nearest adaptive peak, but it need not be the highest peak. The deflections of the evolutionary path caused by the G-matrix may influence which peak is reached but have no partiality for higher peaks. Hence, selection may often act as a constraint that keeps populations from finding better adaptations, and stochastic peak shifts have a potentially important role in allowing populations to explore the broader adaptive landscape. Lande (1985), however, showed that the expected waiting time to attain a successful peak shift scales with the drop in fitness when crossing the valley to the power of N_e, which means that only small populations can cross only shallow valleys in a reasonable time, or to put it bluntly: Significant peak shifts do not occur by drift. Thus populations will typically find themselves trapped on local peaks, and larger evolutionary changes must involve some form of change in the adaptive landscape.

5.7 The Environment in EQG

One attraction of quantitative genetics is the incorporation of environmental sources of individual differences. Despite frequent appeals to high heritability, most variation in quantitative traits tend to be environmental in origin. Hansen and Pélabon (2021) found that the median heritability from 2,536 estimates from wild populations was 31%, meaning that more than 2/3 of the variance is not additive genetic. Even for 1-dimensional morphological measurements, which tend to have the highest heritabilities, only a median 40% of the variance was additive genetic. For life-history traits, which are more susceptible to environmental influences, the median heritability was merely 18%. Some of the nonadditive variance may still be genetic, but environmental sources of variation are clearly important.

The environmental variance can also be broken down into different sources. One important component in many organisms is the effect of parents on their offspring. In mammals, where maternal care and provision are all-important, the maternal component of variance can be substantial (Mousseau and Fox 1998). Parental effects must be considered when estimating genetic variance components and can have consequences for evolution by selection (e.g., Lynch 1987; Kirkpatrick and Lande 1989; Arnold 1994; Wilson et al. 2005; Day and Bonduriansky 2011).

It is convenient to distinguish between micro- and macroenvironmental effects. The former are the effects of many small or unknown sources that must be treated stochastically, as is done with the environmental variance components. Macroenvironmental effects, however, are the result of variation in influential and measurable environmental variables that are somewhat stable for the individual in question. These are often described with reaction norms, which are functions that describe how the expected trait value of a genotype varies with an environmental variable (figure 5.5); for example, body size in relation to food availability. Genetic variation in the slope or shape of reaction norms, that is, genotype-by-environment interaction, sets up the potential for evolutionary changes in plasticity (Via and Lande 1985). Plasticity usually refers to reaction norms that are adapted to create favorable trait values in the particular environment encountered by the individual. The ability to express defense compounds or morphology in the presence of predators or herbivores are standard examples. Not all norms of reaction are adaptive, however; for

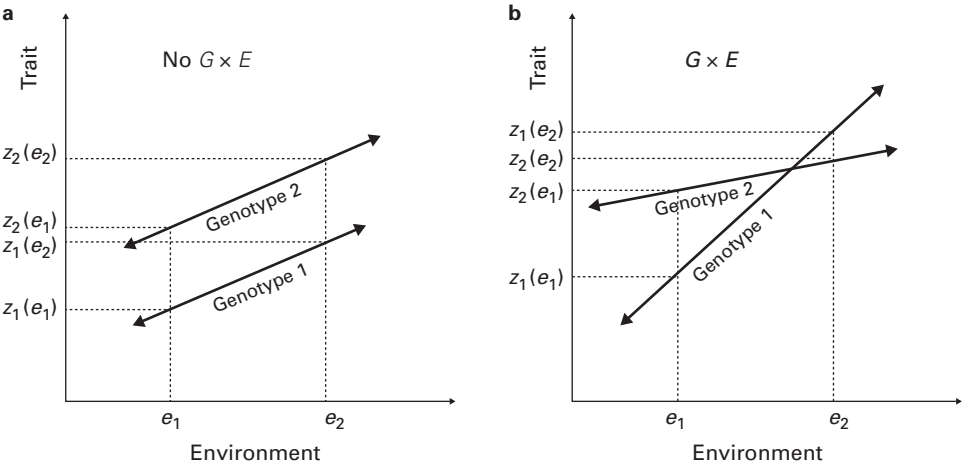


Figure 5.5
Reaction norms: The reaction norms of two genotypes are plotted, and the resulting trait values are given for two states of an environmental variable. Panel A shows a case with no $G \times E$ interaction. The reaction norms have the same shape, so that the phenotypic differences between genotype 1 and genotype 2 are the same in all environments. In panel B, the shapes of the reaction norms differ, so that the difference between the genotypes in environment e_1 is different from the difference in environment e_2, which yields $G \times E$ interaction. With $G \times E$ interaction, the shape of the reaction norm, and thus plasticity, can evolve.

example, growing to small size on limited food is more an outcome of necessary constraints than an adaptive plastic response.

Whether adaptive or not, plasticity and norms of reaction create substantial variation both within and between populations. In a review of reciprocal-transplant experiments, Stamp and Hadfield (2020) found that some 70% of local population differences could be ascribed to plastic responses and only 30% to genetic differences. Plasticity is not just an alternative to genetic adaptation, however, and there has been much discussion on how the two can interact in evolution. Depending on circumstances, plasticity can either facilitate or buffer genetic adaptation (e.g., Paenke et al. 2007).

Traits with lots of genetic variance also tend to have lots of environmental variance. One surprising consequence of this is that heritabilities do no not reflect amount of additive genetic variance (Houle 1992; Hansen et al. 2011; see figure 5.6). The most likely explanation for this puzzling finding is that traits that are generally decanalized or complex with many parts and developmental contingencies tend to be simultaneously sensitive to genetic and environmental perturbations. Cheverud (1988) conjectured that patterns of genetic and environmental correlation are also similar, and based on this, he controversially proposed that the phenotypic P-matrix could be used as substitute for the G-matrix. This idea is approaching a methodological principle in paleobiology and morphometrics, where P-matrices are routinely used in place of G-matrices to study evolvability and constraints (Love et al. 2022). Although we have seen that substituting **P** for **G** cannot be justified by unspecific appeals to "high" heritability, it is an open question when the environmental E-matrix is sufficiently similar to **G** to make **P** a more accurate estimator of the shape of **G** than the often imprecisely estimated **G** itself.

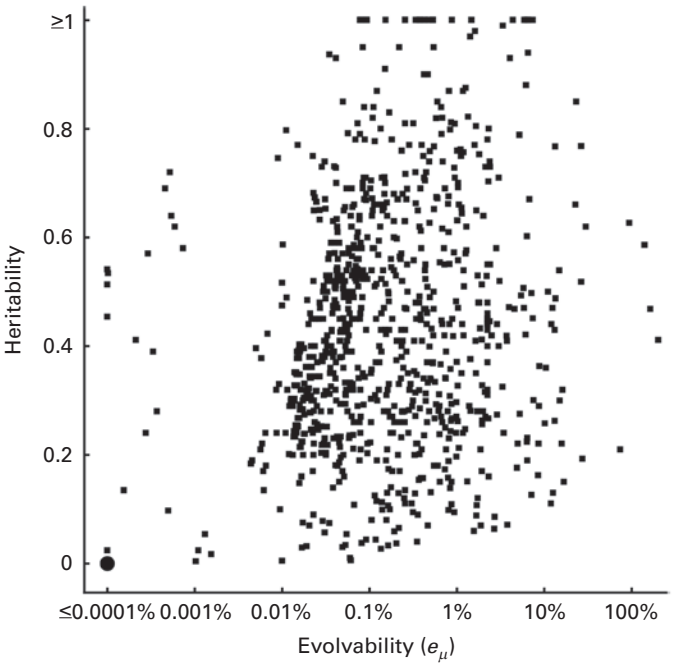


Figure 5.6
Effects of scaling: No relationship between heritability and mean-scaled additive variance (evolvability), as illustrated for 929 1-dimensional morphological measurements. The correlation is 0.04, which means that only a fraction of a percent of the variance in one variable is explained by the other. The big dot in the lower-left corner represents 24 nonpositive estimates. Data from Hansen and Pélabon (2021).

5.8 Issues of Scale and Measurement

In contrast to classical evolutionary genetics, which is concerned with categorical traits on nominal scales, quantitative genetics works with "quantitative" traits on continuous scales. This provides the opportunity for more accurate predictions and tests of the theory, and facilitates better measurement, statistics and quantitative interpretation. Unfortunately, this quantitative promise is often unfulfilled. Evolutionary biologists are accustomed to advanced statistics but are often unaccustomed with quantification. A tradition has developed that allows qualitative interpretations of quantitative results through significance tests and a general neglect of magnitudes, scales and meanings of estimated parameters (Houle et al. 2011; Morrissey 2016).

One manifestation of this tradition is a widespread neglect of units. To see how units work in the focal Lande equation, consider its application to two traits in an imaginary study of sexual selection in a peacock-like bird: display duration, z_1, measured in seconds, and tail area, z_2, measured in square centimeters. Assume a selection gradient is estimated as $\beta = \{0.03\ s^{-1},\ 0.01\ cm^{-2}\}^T$. Note that units of the gradient are inverses of trait units. The partial selection gradient on display duration, $\beta_1 = 0.03\ s^{-1}$, says that if display duration is increased by one second, then fitness (i.e., the fitness component in question) is increased by 3% of its mean. This is because a selection-gradient analysis must use relative fitness, which is fitness measured in units of its own mean. The unit of β_1 is thus fitness per second,

but as fitness is measured on a proportional scale, its units drop out. Similarly, fitness is predicted to increase by 1% per square centimeter increase in tail area.

Technically, the fitness and thus the gradient are per selective episode. This is usually implicit but must be taken into account when making predictions about per generational change, as selective episodes may not be equivalent to generations. Fitnesses and gradients of subsequent selective episodes are multiplicative, whereas those of parallel episodes are additive (weighted with the number of individuals involved).

Assume now that the additive variances of the two traits are $V_{A1} = 5\ s^2$ and $V_{A2} = 300\ cm^4$, and that their covariance is $C_{A12} = 10\ s\ cm^2$. Note that the units of variances are squares of the trait units and that units of covariances are the products of the units of the two traits. Now we can use the Lande equation to predict the responses to selection in the two traits:

$$\Delta\bar{z}_1 = V_{A1}\beta_1 + C_{A12}\beta_2 = 5\,s^2 \cdot 0.03\,s^{-1} + 10\,s\,cm^2 \cdot 0.01\,cm^{-2} = 0.25s,$$
$$\Delta\bar{z}_2 = V_{A2}\beta_2 + C_{A12}\beta_1 = 300\,cm^4 \cdot 0.01\,cm^{-2} + 10\,\mathrm{s}\,cm^2 \cdot 0.03\,s^{-1} = 3.3\,cm^2.$$

We may be interested in which trait evolves the fastest, which trait is most evolvable, and which trait is under strongest direct and indirect selection, but we cannot answer such questions without consideration of scale. The question of whether or how much a change of $3.3\ cm^2$ is larger than $0.25\ s$ is technically meaningless. To make comparisons, we need a common scale.

In EQG, there are three general ways to standardize traits: (1) divide by the population phenotypic standard deviation, (2) divide by the population mean, and (3) use a log scale. I will refer to the first as variance standardization and to the second as mean standardization. Working on a log scale is in practice the same as mean standardization. Both are proportional scales. The difference is that log transformation expresses trait changes in proportion to the trait value, while mean standardization expresses it in proportion to the trait mean. Unless changes are large, numerical and interpretational differences between mean and log scaling are minute.

Variance standardization is most common. It converts additive genetic variances into heritabilities, expresses selection gradients in units of fitness change per standard deviation change in the trait, and gives selection responses and differentials in units of standard deviations. On the assumption that phenotypic standard deviations have the same meaning for different traits, this allows comparison and answers to the above questions. It is problematic, however, because the phenotypic standard deviation is itself intertwined with the entities it is used to scale; additive variances and covariances in particular This creates a correlation between the measure stick and the entities to be measured that can obscure interpretation.

As illustrated in figure 5.6, there is no relationship between variance-standardized additive variances (i.e., heritabilities) and mean-standardized additive variances. This is caused by the correlation between genetic and environmental variances, and probably also by the fact that epistatic variance components scale with powers of the additive variance (Hansen and Wagner 2001), which means that increasing additive variance tends to decrease the proportion of the genetic variance that is additive. So the choice of scale is not trivial, and specifically, heritabilities should not be used to measure evolvability (Hansen et al. 2011).

Houle (1992) was the first to demonstrate problems with variance standardization and to propose mean standardization as an alternative. He showed that life-history traits, which have low heritabilities, in fact tend to have high amounts of additive genetic variation,

whereas morphological traits, which have larger heritabilities, in fact tend to have less additive variation than life-history traits. This finding overturned the hypothesis that stronger selection tends to remove genetic variation from life-history traits and paved the way for the hypothesis that levels of additive variation are more influenced by mutational input. Fitness itself often has high evolvability, but typically very low heritability (Hendry et al. 2018; Hansen and Pélabon 2021).

If we apply mean standardization to the univariate Lande equation, we get

$$\frac{\Delta\bar{z}}{\bar{z}} = \left(\frac{V_A}{\bar{z}^2}\right)(\beta\bar{z}) = e_\mu \beta_\mu,$$

where e_μ is the mean-scaled "evolvability," and β_μ is the mean-scaled selection gradient (Hansen et al. 2003b; Hereford et al. 2004). The mean-scaled selection gradient gives the proportional increase in fitness with a proportional increase in the trait. The mean-scaled evolvability can be interpreted as the predicted proportional increase in the trait mean per generation under unit selection ($\beta_\mu = 1$). Note that mean-scaled evolvabilities are often given as coefficients of additive variation, CV_A, the square root of e_μ, but this is not a good quantitative measure of evolvability, because the response to selection is proportional to trait variance and not to standard deviation.

If mean display duration is $\bar{z}_1 = 10\ s$ and mean tail area is $\bar{z}_2 = 200\ cm^2$, then mean standardizing our example yields

$$\frac{\Delta\bar{z}_1}{\bar{z}_1} = \left(\frac{V_{A1}}{\bar{z}_1^2}\right)(\beta_1\bar{z}_1) + \left(\frac{C_{A12}}{\bar{z}_1\bar{z}_2}\right)(\beta_2\bar{z}_2) = 0.05 \cdot 0.3 + 0.005 \cdot 2 = 0.025,$$

$$\frac{\Delta\bar{z}_2}{\bar{z}_2} = \left(\frac{V_{A2}}{\bar{z}_2^2}\right)(\beta_2\bar{z}_2) + \left(\frac{C_{A12}}{\bar{z}_1\bar{z}_2}\right)(\beta_1\bar{z}_1) = 0.0075 \cdot 2 + 0.005 \cdot 0.3 = 0.017,$$

which predicts that mean display duration will increase by 2.5% per generation, while mean tail area will increase by 1.7%. Hence, on a proportional scale, display duration would evolve faster. Tail area has a stronger effect on fitness, however, because $\beta_{\mu 2} = 2$ means that a 1% increase in tail area would increase fitness by 2% in comparison to $\beta_{\mu 1} = 0.3$, meaning that a 1% increase in display duration would increase fitness by 0.3%. The univariate evolvability of display duration, $e_\mu(z_1) = 0.05$, predicting a 5% increase per generation under unit selection, is much larger than for tail area, however, as $e_\mu(z_2) = 0.0075$ predicts a 0.75% increase in tail area per generation under unit selection. The strong selection on tail area also induces more indirect selection on display duration, contributing to its larger selection response. Although the two traits reinforce each other in this example, we can compute conditional evolvabilities to show that they do not contain much potential for constraining each other. The conditional evolvabilities are $c_\mu(z_1 \mid z_2) = 0.047$ and $c_\mu(z_2 \mid z_1) = 0.0070$, implying a mere 7% reduction in evolvability of each trait if the other was under stabilizing selection.

Note that the selection gradient β measures strength of selection in terms of effect of potential change on fitness, regardless of variation in the trait. The selection differential S, in contrast, depends on trait variation, and if there were a lot of environmental variation in display duration, which would be likely, its selection differential may well be larger than that of tail area. This would not change the response to selection, however, as the environmental variation

would not be transmitted. It would manifest as a lower heritability of display duration. An issue with the breeder's equation is thus that its components, the selection differential and the heritability, tend to be negatively correlated. Counterintuitively, traits with high heritabilities or large selection differentials are not more likely to evolve fast.

Proportional scales provide an intuitive way to quantify evolvability. The number of generations it takes to increase a trait with a factor k is approximately $t_k = \ln(k)/e_\mu\beta_\mu$, and we can take the number of generations to double the trait under unit selection, $t_2 = \ln(2)/e_\mu$, as a convenient quantification of evolutionary potential. Applying this to estimates of evolvability taken from the recent review of Hansen and Pélabon (2021) shows that quantitative traits usually have large potentials for change on macroevolutionary time scales. Although the median evolvability for 1-dimensional morphological traits of $e_\mu \approx 0.1\%$ sounds tiny, it allows a trait to double in the geological eyeblink of $t_2 \approx 700$ generations under unit selection. The median evolvability of fitness is $e_\mu \approx 1.3\%$, and because fitness is under unit directional selection by definition, this number implies that selection would continuously double genetic fitness each 50 generations if no other forces were involved.

Thoughtless standardizations have caused much interpretational damage in EQG and related fields. The misapplication of heritability as a measure of evolutionary potential is the most obvious example, but variance standardization has also wreaked havoc on many studies of selection and rates of evolution. It is particularly dangerous in multivariate studies, in which some form of standardization is usually required to compare traits. In studies based on P-matrices, variance standardization will convert P-matrices to correlation matrices; all information about trait variation will be lost; and any inference about modularity, integration, and evolutionary potential will be dubious.

Mean standardization and log scales largely avoid the rubber-scale problems of variance standardization, but some traits are not meaningful or need careful interpretation on proportional scales (e.g., Hereford et al. 2004; Stinchcombe 2005; Hansen et al. 2011; Houle et al. 2011; Matsumura et al. 2012; Opedal et al. 2017; Pélabon et al. 2020). As much as possible, interpretation should involve original scales, biological information, and theoretical context. Even if a quarter-second change in display duration cannot be directly compared to a square-centimeter change in tail area, these figures may have biological meaning to a researcher familiar with the organism and relevant theory. Results should be back-transformed and evaluated on original scales. This is particularly important when arbitrary nonlinear transformations have been applied for statistical reasons.

5.9 Quantitative Genetics and the Genotype-Phenotype Map

The statistical representation of variation, inheritance, and selection is well suited to describe evolution with simple operational models. Its disadvantage is that underlying biological complexities are hidden, and sometimes this matters for understanding evolution and particularly for understanding constraints. The G- (and M-) matrix only captures linear and often temporary constraints. To understand evolutionary potential for larger changes, the underlying biological causality needs be considered. With the application of quantitative genetics to evolutionary questions, this became increasingly apparent, and the effects of underlying development, physiology and genetic architecture on quantitative genetic variation are increasingly taken into consideration.

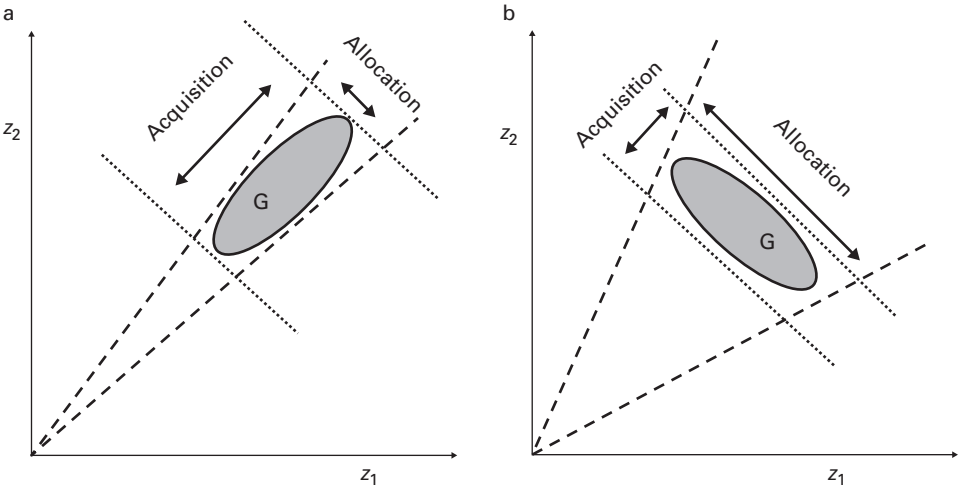


Figure 5.7
Effects of pleiotropy on the G-matrix. Two traits, z_1 and z_2, are determined by two genetically determined processes controlling how much resources are acquired and allocated to the growth of the traits. In (a) there is more variation in acquistion, causing positive covariance between the traits. In (b) there is more variation in the allocation, causing negative covariance between the traits.

One example, illustrated in figure 5.7, shows how two traits that are integrated on the biological level can still display all kinds of covariance structures. The traits are assumed to grow together on a shared resource that is then allocated between them, and their covariance is determined by how much variation there is in the shared resource and in the mechanism of allocation. This model has been used to illustrate problems with using the G- (or P-) matrix to study trade-offs or constraints between traits (e.g., Riska 1986; Houle 1991; Fry 1993). Even if there is a biological trade-off, covariances may still be positive or absent. More generally, it illustrates how a mixture of positive and negative pleiotropy can generate different patterns of covariance between traits.

Although pleiotropy (allelic differences that affect several traits) will usually be the main source of genetic covariance between traits, genetic covariance can also arise through linkage disequilibrium (covariance in occurrence) between alleles affecting the respective traits. Covariance through linkage disequilibrium can arise from genetic drift or nonrandom mating. It can also be built by correlational selection, but because it is broken down by recombination, it is unlikely to be substantial for polygenic traits with many unlinked sets of loci (e.g., Lande 1984). Gains from selection on variation due to linkage disequilibrium will be transient, but they may still be important if there are special mechanisms maintaining the disequilibrium. One example arises in the Fisher runaway model of sexual selection, when assortative mating between individuals with a particular trait and individuals with a mating preference for that trait creates linkage disequilibrium between alleles affecting the trait and alleles affecting the preference. Subsequent selection on the trait caused by the mating preference in the population will then also cause indirect selection on the preference genes. Hence, the preference indirectly selects itself, causing a potential for runaway evolution.

The main victim of the statistical representation of the genotype-phenotype map has been epistasis. There is a widespread misconception in quantitative genetics that epistasis

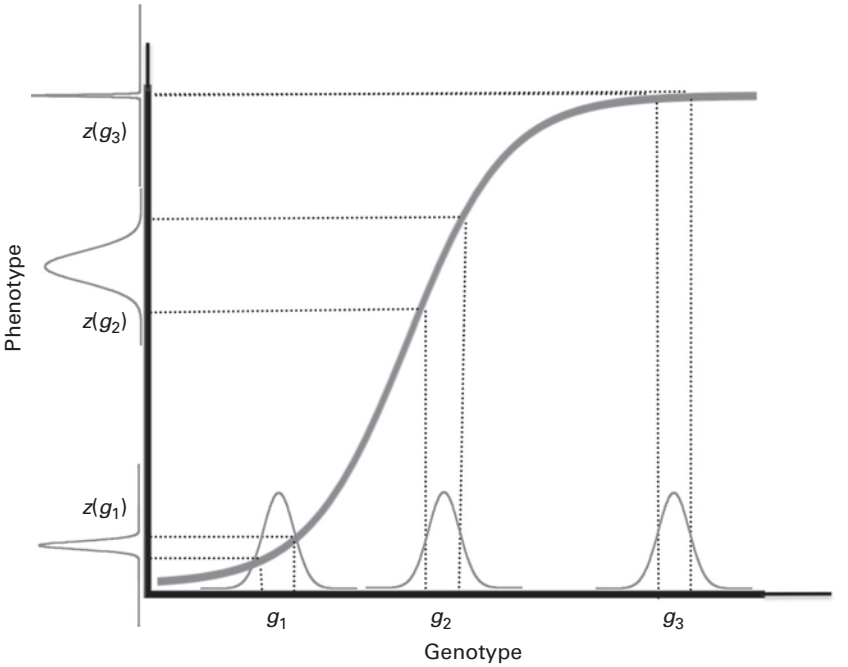


Figure 5.8
The genotype-phenotype map and the evolution of evolvability: Three different genetic backgrounds, g_1, g_2, and g_3, map the same amount of molecular variation to different levels of genetic variation in the phenotype. From g_1 to g_2 positive directional epistasis increases variation, and from g_2 to g_3 negative directional epistasis decreases variation. Note that there is negative directional epistasis in both directions from g_2. The phenotype is constrained from evolving beyond the range from $z(g_1)$ to $z(g_3)$. Based on Hansen (2015).

has only transient effects on evolution by selection and can be ignored. This conclusion is based on the finding that selection on epistatic variance only leads to changes in linkage disequilibrium and not to allele-frequency changes. Creating epistatic variance is not the only influence of epistasis, however. Biological epistasis also changes the biological (and thus statistical) effects of alleles, which may change the additive variance and thus the response to selection. Such changes are not transient and may be substantial (Hansen 2013). A key point is that permanent effects depend on systematic patterns of epistasis (Carter et al. 2005). Systematic positive epistasis in which allele substitutions with positive effects on the trait also tend to increase the effects of other allele substitutions on the trait will increase the additive variance and accelerate the response of a positively selected trait (figure 5.8). Systematic negative epistasis will have the opposite effects and can lead to canalization and evolutionary standstill (Hansen et al. 2006; Le Rouzic and Alvarez-Castro 2016). Nondirectional epistasis, in which positive and negative effects balance, will have no immediate effect on the response to selection. The statistical description of epistasis is blind to biological patterns in the interactions in the sense that the epistatic variance components are the same, regardless of presence or sign of directionality. EQG thus lacked operational tools to quantify relevant effects of epistasis, and some prominent researchers confused the absence of epistasis from quantitative-genetics theory with evidence against its importance. This gap is being ameliorated by consideration of explicit, theoretical or empirical genotype-phenotype maps (e.g., Cheverud and Routman 1995; Hansen and

Wagner 2001; Alvarez-Castro and Carlborg 2007; Le Rouzic 2014; Morrissey 2014, 2015; Hansen 2015; Alvarez-Castro 2016; Milocco and Salazar-Ciudad 2020), but empirical research on relevant patterns of epistasis is still scarce.

The emergence of genomics and molecular methods has provided many new insights into the genetic architecture of quantitative traits. Marker-based QTL and GWAS studies not only identify genes affecting traits but are also increasingly used to illuminate the general structure of the genotype-phenotype map. Here I only highlight one recent development: The increasing realization that many traits are influenced by segregating variation at thousands of locations in the genome (e.g., Boyle et al. 2017; Pitchers et al. 2019; Jakobson and Jarosz 2020). This conclusion contrasts with the picture provided by the first molecular analyses of quantitative traits in the 1990s, which often identified one or a few genes with large effects. These were mostly statistical artifacts, and even later studies finding dozens or hundreds of "significant" genes could usually not explain more than a fraction of the genetic variance. This became known as the problem of the "missing heritability." Subsequent studies with larger samples and better methods have traced the missing genetic variance to a huge number of segregating variants with very small effects. This "omnigenic" model (Boyle et al. 2017; Liu et al. 2019) implies that every gene in the genome has potential and often real effects on every quantitative trait. The likely reason is that a multitude of convoluted, indirect pathways exist for genes to affect traits, which again implies universal pleiotropy and a huge potential for the influence of epistatic interactions and for the generation of novel trait values by recombination (Hansen and Pélabon 2021). While there must be variation in the degree of directness of how genes influence a trait, it is debatable whether this model implies a distinction between core and peripheral genes as suggested by Liu et al. (2019).

5.10 Conclusion: Explaining Macroevolution?

The EQG research program was motivated by Lande's audacious hypothesis that simple evolutionary predictions based on a constant G-matrix could illuminate macroevolutionary change. Today there are vibrant research fields that use the EQG framework to understand short- and long-term evolution in the context of natural and sexual selection, life-history theory, behavior, and morphology (e.g., Boake 1994; Schluter 2000; Arnold et al. 2001; Pigliucci and Preston 2004; Roff 2007; Polly 2008; Svensson and Calsbeek 2012; Arnold 2014; Charmantier et al. 2014; Hendry 2017). The framework is frequently used to interpret long-term evolution in paleontological time series or comparative data. While the literal extrapolation of simple EQG models beyond a handful of generations may seem futile, as neither selection nor evolvability are likely to remain constant when the phenotype is changing, the EQG approach has still contributed many theoretical and empirical insights that must be part of the foundation for quantitative analysis of both micro- and macroevolution. These include:

1. Genetic architectures are highly polygenic and pleiotropic.
2. Mutational input to quantitative traits is high and explains variation in variation.
3. Short-term evolvability depends on additive variance.
4. Heritability is not evolvability.
5. Evolvability of low-dimensional traits is high from a macroevolutionary perspective.

6. Selection is strong, directional, and fluctuating on short time scales.
7. Selection is stabilizing over larger phenotypic and temporal ranges.
8. Indirect selection is important.
9. Neutral evolution is fast.
10. Peak shifts by drift do not occur.
11. Plasticity is important.
12. The evolution of evolvability (additive and mutational variance) depends on epistasis.
13. Microevolution is rapid and fluctuating but usually stationary.

An intuitive inference is that macroevolutionary dynamics should be determined by changes in the adaptive landscape and that genetic constraints are unimportant. At least the latter part of this inference is premature, however. There are indications that the structure of the G-matrix is related to rates of evolution and among-species variation even on million-year time scales (Voje et al., chapter 14),[1] and even if high evolvabilities allow rapid microevolution, we lack an understanding of how far traits can be extended without inducing pleiotropic or epistatic constraints.

Pleiotropic constraints arise when selection on a focal trait due to pleiotropy induces indirect selection on the genetic basis of other traits. Such changes are unlikely to be favorable and may cause counteracting selection to repair the damage. This effect is what is captured by conditional evolvability, but the measures I have discussed only quantify the effects of specific macroscopic traits, while we would like to know the effects from the entire genome. Universal pleiotropy predicts that selection on a trait would induce multitudes of minor changes in the genetic basis of essentially all other traits, and the consequences of this remains to be worked out.

Epistatic constraints arise when selection reduces genetic variance through negative directional epistasis. A genotype-phenotype map as shown in figure 5.8 would induce negative epistasis in both directions from the middle, and if the trait is selected toward the edges, it would become canalized and lose evolvability. Whether such structuring of genotype-phenotype maps is common is unknown, but it is a potential explanation for stationarity of microevolutionary change.

In conclusion, EQG provides a powerful toolbox for quantitative analysis of microevolution, but fundamental misconceptions of scale and the lack of a distinction between biological and statistical effects have hampered empirical and theoretical interpretation. When these problems are recognized, EQG should be a pillar for interpreting both micro- and macroevolution.

Acknowledgments

Thanks to the members of the evolvability group at the Centre for Advanced Study in Oslo and to members of the evolvability journal club for discussions and comments. David Houle and Christophe Pélabon made helpful comments on the manuscript.

1. References to chapter numbers in the text are to chapters in this volume.

References

Agrawal, A. F., and J. R. Stinchcombe. 2009. How much do genetic covariances alter the rate of adaptation? *Proceedings of the Royal Society B* 276: 1182–1191.

Alvarez-Castro, J. M. 2016. Mapping genetic architectures: A shift in evolutionary quantitative genetics. In *The Encyclopedia of Evolutionary Biology,* vol. 2, edited by R. M. Kliman, 127–135. Oxford: Academic Press.

Alvarez-Castro, J. M., and O. Carlborg. 2007. A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis. *Genetics* 176: 1151–1167.

Arnold, S. J. 1983. Morphology, performance and fitness. *American Zoologist* 23: 347–361.

Arnold, S. J. 1992. Constraints on phenotypic evolution. *American Naturalist* 140: S85–S107.

Arnold, S. J. 1994. Multivariate inheritance and evolution: A review of concepts. In *Quantitative Genetic Studies of Behavioral Evolution,* edited by C. R. B. Boake, 17–48. Chicago: University of Chicago Press.

Arnold, S. J. 2014. Phenotypic evolution: The ongoing synthesis. *American Naturalist* 183: 729–746.

Arnold, S. J., and M. J. Wade. 1984a. On the measurement of natural and sexual selection: Theory. *Evolution* 38: 709–719.

Arnold, S. J., and M. J. Wade. 1984b. On the measurement of natural and sexual selection: Applications. *Evolution* 38: 720–734.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112/113: 9–32.

Arnold, S. J., R. Bürger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones. 2008. Understanding the evolution and stability of the G-matrix. *Evolution* 62: 2451–2461.

Beatty J. 2016. The creativity of natural selection? Part I: Darwin, Darwinism, and the mutationists. *Journal of the History of Biology* 49: 659–684.

Berg, R. L. 1959. A general evolutionary principle underlying the origin of developmental homeostasis. *American Naturalist* 93: 103–105.

Blows, M. W. 2007. A tale of two matrices: Multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* 20: 1–8.

Boake, C. R. B. (Editor). 1994. *Quantitative Genetic Studies of Behavioral Evolution.* Chicago: University of Chicago Press.

Bolstad, G. H., T. F. Hansen, C. Pélabon, M. Falahati-Anbaran, R. Pérez-Barrales, and W. S. Armbruster. 2014. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. *Philosophical Transactions Royal Society B* 369: 20130255.

Boyle, E. A., Y. I. Li, and J. K. Pritchard. 2017. An expanded view of complex traits: From polygenic to omnigenic. *Cell* 169: 1177–1186.

Bulmer M. 2003. *Francis Galton: Pioneer of Heredity and Biometry.* London: The Johns Hopkins University Press.

Carter, A. J. R., J. Hermisson, and T. F. Hansen. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theoretical Population Biology* 68: 179–196.

Charmantier, A., D. Garant, and L. E. B. Kruuk. (Editors). 2014. *Quantitative Genetics in the Wild.* Oxford: Oxford University Press.

Cheng, C., and D. Houle. 2020. Predicting multivariate responses of sexual dimorphism to direct and indirect selection. *American Naturalist* 196: 391–405.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. *Journal of Theoretical Biology* 110: 155–171.

Cheverud, J. M. 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42: 958–968.

Cheverud, J. M., and E. J. Routman. 1995. Epistasis and its contribution to genetic variance components. *Genetics* 139: 1455–1461.

Chevin, L. M. 2013. Genetic constraints on adaptation to a changing environment. *Evolution* 67: 708–721.

Day, T., and R. Bonduriansky. 2011. A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *American Naturalist* 178: E18–E36.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* 3: 399–433.

Fry, J. D. 1993. The "general vigor" problem: Can antagonistic pleiotropy be detected when genetic covariances are positive? *Evolution* 47: 327–333.

Grabowski, M., and A. Porto. 2017. How many more? Sample size determination in studies of morphological integration and evolvability. *Methods in Ecology and Evolution* 8: 592–603.

Gudbjartsson, D. F., G. B. Walters, G. Thorleifson, et al. 2008. Many sequence variants affecting diversity of adult human height. *Nature Genetics* 40: 609–615.

Halligan, D. L., and P. D. Keightley. 2009. Spontaneous mutation accumulation studies in evolutionary genetics. *AREES* 40: 151–172.

Hansen, T. F. 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. *Biosystems* 69: 83–94.

Hansen, T. F. 2013. Why epistasis is important for selection and adaptation. *Evolution* 67: 3501–3511.

Hansen, T. F. 2015. Measuring gene interactions. In *Epistasis: Methods and Protocols,* edited by J. H. Moore, and S. M. Williams, 115–143. New York: Humana Press.

Hansen, T. F. 2017. On the definition and measurement of fitness in finite populations. *Journal of Theoretical Biology* 419: 36–43.

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* 21: 1201–1219.

Hansen, T. F., and C. Pélabon. 2021. Evolvability: A quantitative-genetics perspective. *AREES* 52: 153–175.

Hansen, T. F., and K. L. Voje. 2011. Deviation from the line of least resistance does not exclude genetic constraints: A comment on Berner et al. (2010). *Evolution* 65: 1821–1822.

Hansen, T. F., and G. P. Wagner. 2001. Modeling genetic architecture: A multilinear model of gene interaction. *Theoretical Population Biology* 59: 61–86.

Hansen, T. F., W. S. Armbruster, M. L. Carlson, and C. Pélabon. 2003a. Evolvability and genetic constraint in *Dalechampia* blossoms: Genetic correlations and conditional evolvability. *Journal of Experimental Zoology B* 296: 23–39.

Hansen, T. F., C. Pélabon, W. S. Armbruster, and M. L. Carlson. 2003b. Evolvability and genetic constraint in *Dalechampia* blossoms: Components of variance and measures of evolvability. *Journal of Evolutionary Biology* 16: 754–765.

Hansen, T. F., J. M. Alvarez-Castro, A. J. R. Carter, J. Hermisson, and G. P. Wagner. 2006. Evolution of genetic architecture under directional selection. *Evolution* 60: 1523–1536.

Hansen, T. F., C. Pélabon, and D. Houle. 2011. Heritability is not evolvability. *Evolutionary Biology* 38: 258–277.

Hansen, T. F., T. M. Solvin, and M. Pavličev. 2019. Predicting evolutionary potential: A numerical test of evolvability measures. *Evolution* 73: 689–703.

Hendry, A. P. 2017. *Eco-evolutionary Dynamics.* Princeton, NJ: Princeton University Press.

Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The contemporary evolution of fitness. *AREES* 49: 457–476.

Hereford, J., T. F. Hansen, and D. Houle. 2004. Comparing strengths of directional selection: How strong is strong? *Evolution*: 58: 2133–2143.

Hermisson, J., T. F. Hansen, and G. P. Wagner. 2003. Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. *American Naturalist* 161: 708–734.

Hine, E., S. F. Chenoweth, H. D. Rundle, and M. W. Blows. 2009. Characterizing the evolution of genetic variance using genetic covariance tensors. *Philosophical Transactions of the Royal Society B* 364: 1567–1578.

Hohenlohe, P. A., and S. A. Arnold. 2008. MiPoD: A hypothesis-testing framework for microevolutionary inference from patterns of divergence. *American Naturalist* 171: 366–385.

Houle, D. 1991. Genetic covariance of fitness correlates: What genetic correlations are made of and why it matters. *Evolution* 45: 630–648.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130: 195–204.

Houle, D. 1998. How should we explain variation in the genetic variance of traits? *Genetica* 102/103: 241–253.

Houle, D., and A. Kondrashov. 2006. Mutation. In *Evolutionary Genetics: Concepts and Case Studies,* edited by C. W. Fox and J. B. Wolf, 32–48. Oxford: Oxford University Press.

Houle D., C. Pélabon, G. P. Wagner, and T. F. Hansen. 2011. Measurement and meaning in biology. *Quarterly Review of Biology* 86: 3–34.

Jakobson, C. M., and D. F. Jarosz. 2020. What has a century of quantitative genetics taught us about nature's genetic toolkit? *Annual Review of Genetics* 54: 439–464.

Jensen, H., M. Szulkin, and J. Slate. 2014. Molecular quantitative genetics. In *Quantitative Genetics in the Wild,* edited by A. Charmantier, D. Garant, and L. E. B. Kruuk, 209–227. Oxford: Oxford University Press.

Kerr, B., and P. Godfrey-Smith. 2009. Generalization of the Price equation for evolutionary change. *Evolution* 63: 531–536.

Kirkpatrick, M. 2009. Patterns of quantitative genetic variation in multiple dimensions. *Genetica* 136: 271–284.

Kirkpatrick, M., and R. Lande. 1989. The evolution of maternal characters. *Evolution* 43: 485–503.

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

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

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. 1984. The genetic correlation between characters maintained by selection, linkage and inbreeding. *Genetics Research* 44: 309–320.

Lande, R. 1985. Expected time for random genetic drift of a population between stable phenotypic states. *PNAS* 82: 7641–7645.

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

Le Rouzic, A. 2014. Estimating directional epistasis. *Frontiers in Genetics* 5: 198.

Le Rouzic, A., and J. M. Alvarez-Castro. 2016. Epistasis-induced evolutionary plateaus in selection responses. *American Naturalist* 188: E134–E150.

Liu, X., Y. I. Li, and J. K. Pritchard. 2019. Trans effects on gene expression can drive omnigenic inheritance. *Cell* 177: 1022–1034.

Love, A. C., M. Grabowski, D. Houle, et al. 2022. Evolvability in the fossil record. *Paleobiology* 48: 186–209.

Lynch, M. 1987. Evolution of intrafamilial interactions. *PNAS* 84: 8507–8511.

Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *American Naturalist* 136: 727–741.

Lynch, M. 2007. *The Origins of Genome Architecture.* Sunderland, MA: Sinauer.

Lynch, M., and B. Walsh. 1998. *Genetics and Analysis of Quantitative Characters*. Sunderland, MA: Sinauer.

Majerus, M. E. N. 1998. *Melanism: Evolution in Action*. Oxford: Oxford University Press.

Marquez, E. J. 2008. A statistical framework for testing modularity in multidimensional data. *Evolution* 62: 2688–2708.

Marroig, G., L. T. Shirai, A. Porto, F. B. de Oliveria, and V. De Conto. 2009. The evolution of modularity in the mammalian skull II: Evolutionary consequences. *Evolutionary Biology* 36: 136–148.

Matsumura, S., R. Arlinghaus, and U. Dieckmann. 2012. Standardizing selection strengths to study selection in the wild: A critical comparison and suggestions for the future. *BioScience* 62: 1039–1054.

McGlothlin, J. W., M. E. Kobiela, H. V. Wright, D. L. Mahler, J. J. Kolbe, J. B. Losos, and E. D. III Brodie. 2018. Adaptive radiation along a deeply conserved genetic line of least resistance in *Anolis* lizards. *Evolution Letters* 2–4: 310–322.

Melo, D., A. Porto, J. M. Cheverud, and G. Marroig. 2016. Modularity: Genes, development, and evolution. *AREES* 47: 463–486.

Milocco, L., and I. Salazar-Ciudad. 2020. Is evolution predictable? Quantifying the accuracy of quantitative genetics predictions for complex genotype-phenotype maps. *Evolution* 74: 230–244.

Mitteroecker, P. 2009. The developmental basis of variational modularity: Insights from quantitative genetics, morphometrics, and developmental biology. *Evolutionary Biology* 36: 377–385.

Morrissey, M. B. 2014. Selection and evolution of causally covarying traits. *Evolution* 68: 1748–1761.

Morrissey, M. B. 2015. Evolutionary quantitative genetics of non-linear developmental systems. *Evolution* 69: 2050–2066.

Morrissey, M. B. 2016. Meta-analysis of magnitudes, differences, and variation in evolutionary parameters. *Journal of Evolutionary Biology* 29: 1882–1904.

Morrissey, M. B., L. E. B. Kruuk, and A. J. Wilson. 2010. The danger of applying the breeder's equation in observational studies of natural populations. *Journal of Evolutionary Biology* 23: 2277–2288.

Morrissey, M. B., P. de Villemereull, B. Doligez, and O. Gimenez. 2014. Bayesian approaches to the quantitative genetic analysis of natural populations. In *Quantitative Genetics in the Wild,* edited by A. Charmantier, D. Garant, and L. E. B. Kruuk, 228–253. Oxford: Oxford University Press.

Mousseau, T. A., and C. W. Fox. (Editors). 1998. *Maternal Effects as Adaptations.* Oxford: Oxford University Press.

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

Opedal, Ø. H., G. H. Bolstad, T. F. Hansen, W. S. Armbruster, and C. Pélabon. 2017. The evolvability of herkogamy: Quantifying the evolutionary potential of a composite trait. *Evolution* 71: 1572–1586.

Otsuka, J. 2019. Ontology, causality, and methodology of evolutionary research programs. In *Evolutionary Causation: Biological and Philosophical Reflections,* edited by T. Uller and K. N. Laland, 247–264. The Vienna Series in Theoretical Biology. Cambridge, MA: MIT Press.

Paenke, I., B. Sendhoff, and T. J. Kawecki. 2007. Influence of plasticity and learning on evolution under directional selection. *American Naturalist* 170: E47–E58.

Pélabon, C., C. H. Hilde, S. Einum, and M. Gamelon. 2020. On the use of the coefficient of variation to quantify and compare trait variation. *Evolution Letters* 4: 180–188.

Pigliucci, M., and K. Preston. (Editors). 2004. *Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes.* Oxford: Oxford University Press.

Pitchers, W., J. Nye, E. J. Marquez, A. Kowlaski, I. Dworkin, and D. Houle. 2019. A multivariate genome-wide association study of wing shape in *Drosophila melanogaster. Genetics* 211: 1429–1447.

Polly, P. D. 2008. Developmental dynamics and G-matrices: Can morphometric spaces be used to model phenotypic evolution? *Evolutionary Biology* 35: 83–96.

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

Provine, W. B. 1971. *The Origins of Theoretical Population Genetics.* Chicago: University of Chicago Press.

Reid, J. M., and R. J. Sardell. 2011. Indirect selection on female extra-pair reproduction? Comparing the additive genetic value of maternal half-sib extra-pair and within-pair offspring. *Philosophical Transactions of the Royal Society B* 279: 1700–1708.

Riedl, R. J. 1977. A systems-analytical approach to macro-evolutionary phenomena. *Quarterly Review of Biology* 52: 351–370.

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

Roff, D. A. 2007. A centennial celebration for quantitative genetics. *Evolution* 61: 1017–1032.

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

Schluter, D. 2000. *The Ecology of Adaptive Radiation.* Oxford: Oxford University Press.

Stamp, M. A., and J. D. Hadfield. 2020. The relative importance of plasticity versus genetic differentiation in explaining between population differences; a meta-analysis. *Ecology Letters* 23: 1432–1441.

Stinchcombe, J. R. 2005. Measuring natural selection on proportional traits: Comparisons of three types of selection estimates for resistance and susceptibility to herbivore damage. *Evolutionary Ecology* 19: 363–373.

Stinchcombe, J. R., A. F. Agrawal, P. A. Hohenlohe, S. J. Arnold, and M. W. Blows. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: Double or nothing? *Evolution* 62: 2435–2440.

Svensson, E. I., and R. Calsbeek. (Editors). 2012. *The Adaptive Landscape in Evolutionary Biology.* Oxford: Oxford University Press.

Sztepanacz, J. L., and D. Houle. 2019. Cross-sex genetic covariances limit the evolvability of wing-shape within and among species of *Drosophila. Evolution* 73: 1617–1633.

Turelli, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's Zeta meets the abdominal bristle. *Theoretical Population Biology* 25: 138–193.

Via, S., and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39: 505–522.

Wade, M. J., and S. Kalisz. 1990. The causes of natural selection. *Evolution* 44: 1947–1955.

Wagner, G. P. 2014. *Homology, Genes, and Evolutionary Innovation.* Princeton, NJ: Princeton University Press.

Wagner, G. P., and L. Altenberg. 1996. Complex adaptations and the evolution of evolvability. *Evolution* 50: 967–976.

Walsh, B., and M. Lynch. 2018. *Evolution and Selection of Quantitative Traits.* Oxford: Oxford University Press.

Wilson, A. J., D. W. Coltman, J. M. Pemberton, A. D. J. Overall, K. A. Byrne, and L. E. B. Kruuk. 2005. Maternal genetic effects set the potential for evolution in a free-living vertebrate population. *Journal of Evolutionary Biology* 18: 405–414.

Wilson, A. J., D. Reale, M. N. Clements, et al. 2010. An ecologist's guide to the animal model. *Journal of Animal Ecology* 79: 13–26.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Times New Roman by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Hansen, Thomas F., editor.
Title: Evolvability : a unifying concept in evolutionary biology? / edited by Thomas F. Hansen, David Houle, Mihaela Pavličev, and Christophe Pélabon.
Description: Cambridge, Massachusetts : The MIT Press, [2023] | Series: Vienna series in theoretical biology | Includes bibliographical references and index.
Identifiers: LCCN 2022038288 (print) | LCCN 2022038289 (ebook) | ISBN 9780262545624 (paperback) | ISBN 9780262374705 (epub) | ISBN 9780262374699 (pdf)
Subjects: LCSH: Evolution (Biology)—Philosophy.
Classification: LCC QH360.5 .E99 2023 (print) | LCC QH360.5 (ebook) | DDC 576.801—dc23/eng/20220920
LC record available at https://lccn.loc.gov/2022038288
LC ebook record available at https://lccn.loc.gov/2022038289